Selected Tracking and Fusion Applications for the Defence and Security Domain

Wolfgang Koch

Fraunhofer FKIE
Neuenahrer Strasse 20
D 53343 Wachtberg, Germany
Tel +49 228 9435 373
Fax +49 228 9435 685

email w.koch@ieee.org

Abstract

Sensor Data Fusion is the process of combining incomplete and imperfect pieces of mutually complementary sensor information in such a way that a better understanding of an underlying real-world phenomenon is achieved. Typically, this insight is either unobtainable otherwise or a fusion result exceeds what can be produced from a single sensor output in accuracy, reliability, or cost. Appropriate collection, registration and alignment, stochastic filtering, logical analysis, space-time integration, exploitation of redundancies, quantitative evaluation, and appropriate display are part of Sensor Data Fusion as well as the integration of related context information.

Sensor Data Fusion, as an information technology as well as a branch of engineering science and informatics, is discussed in an introductory part, put into a more general context, and related to information systems. Basic elements and concepts are introduced. Selected applications are discussed in the subsequent sections, where specific problems of Sensor Data Fusion are highlighted. The material discussed in the individual sections is collected from journal publications by the author.

Contents

1 Introduction **3**
1.1 Situation Pictures . . . 3
1.2 Origins of Modern Development . . . 3
1.3 General Technological Prerequisites . . . 4
1.4 Relation to Information Systems . . . 5

2 Discussion of a Characteristic Example **6**
2.1 From Imperfect Data to Situation Pictures . . . 7
2.2 Aspects in Situation Picture Production . . . 8
2.3 Remarks on the Methods Used . . . 9
2.4 A Schematic Sensor Data Fusion System . . . 9
2.4.1 Sensor Systems . . . 10
2.4.2 Interoperability . . . 10
2.4.3 Fusion Process . . . 10
2.4.4 Human-Machine Interface . . . 11
2.5 On Measuring Fusion Performance . . . 11
2.6 Tracking-derived Situation Elements . . . 12
2.6.1 Inferences based on Retrodicted Tracks . . . 12
2.6.2 Inferences based on Multiple Target Tracking . . . 12
2.7 Selected Issues in Anomaly Detection . . . 13
2.7.1 Integration of Planning Information . . . 14
2.7.2 Detecting Regularity Pattern Violation . . . 14
2.7.3 Tracking-derived Regularity Patterns . . . 15

3 Integration of Advanced Sensor Properties **15**
3.1 Finite Sensor Resolution . . . 15
3.1.1 A Radar Resolution Model . . . 16
3.1.2 Resolution-specific Likelihood . . . 18
3.1.3 A Formation Tracking Example . . . 20
3.1.4 Resolution: Summary of Results . . . 22
3.2 Tracking in Presence of Main-lobe Jamming . . . 23
3.2.1 Modeling the Jammer Notch . . . 23
3.2.2 Tracking Filters Alternatives . . . 25
3.2.3 Selected Simulation Results . . . 26
3.3 Negative Sensor Information . . . 27
3.3.1 A Ubiquitous Notion . . . 27
3.3.2 Lessons Learned from Examples . . . 27

4 Integration of Advanced Object Properties **28**
4.1 Extended Object Tracking . . . 29
4.1.1 Generalized Formalism . . . 29
4.1.2 Extended Object Prediction . . . 31
4.1.3 Extended Object Filtering . . . 33
4.1.4 Extended Object Kinematics . . . 34
4.1.5 Selected Simulation Results . . . 35
4.1.6 Summary of Results . . . 38
4.2 Classification with Chemical Sensors . . . 40

1 Introduction

Sensor data fusion is an omnipresent phenomenon that existed prior to its technological realization or the scientific reflection on it. In fact, all living creatures, including human beings, by nature or intuitively perform sensor data fusion. Each in their own way, they combine or "fuse" sensations provided by different and mutually complementary sense organs with knowledge learned from previous experiences and communications from other creatures. As a result, they produce a "mental picture" of their individual environment, the basis of behaving appropriately in their struggle to avoid harm or successfully reach a particular goal in a given situation.

1.1 Situation Pictures

As a sophisticated technology with significant economic and defence implications as well as a branch of engineering science and applied informatics, modern sensor data fusion aims at automating this capability of combining complementary pieces of information. Sensor data fusion thus produces a "situation picture", a reconstruction of an underlying "real situation", which is made possible by efficiently implemented mathematical algorithms exploiting even imperfect data and enhanced by new information sources. Emphasis is not only placed on advanced sensor systems, technical equivalents of sense organs, but also on spatially distributed networks of homogeneous or heterogeneous sensors on stationary or moving platforms and on the integration of data bases storing large amounts of quantitative context knowledge. The suite of information sources to be fused is completed by the interaction with human beings, which makes their own observations and particular expertise accessible.

The information to be fused may comprise a large variety of attributes, characterized, for example, by sensor ranges from less than a meter to hundreds of kilometers, by time scales ranging from less than second to a few days, by nearly stationary or rapidly changing scenarios, by actors behaving cooperatively, incooperatively, or even hostile, by high precision measurements or sensor data of poor quality.

Sensor data fusion systems emerging from this branch of technology have in effect the character of "cognitive tools", which enhance the perceptive faculties of human beings in the same way conventional tools enhance their physical strength. In this type of interactive assistance system, the strengths of automated data processing (dealing with mass data, fast calculation, large memory, precision, reliability, robustness etc.) are put into service for the human beings involved. Automated sensor data fusion actually enables them to bring their characteristically "human" strengths into play, such as qualitatively correct over-all judgment, expert knowledge and experience, intuition and creativity, i.e. their "natural intelligence" that cannot be substituted by automated systems in the foreseeable future. The user requirements to be fulfilled in a particular application have a strong impact on the actual fusion system design.

1.2 Origins of Modern Development

Sensor data fusion systems have been developed primarily for applications, where a particular need for support systems of this type exists, for example in time-critical situations or in situations with a high decision risk, where human deficiencies must be complemented by automatically or interactively working data fusion techniques. Examples are fusion tools for compensating decreasing attention in routine and mass situations, for focusing attention on anomalous or rare events, or complementing limited memory, reaction, and combination capabilities of human beings. In addition to the advantages of reducing the human workload in routine or mass tasks by exploiting large data streams quickly, precisely, and comprehensively, fusion of mutually complementary information sources typically produces qualitatively new and important knowledge that otherwise would remain unrevealed.

The demands for developing such support systems are particularly pressing in defence and security applications, such as surveillance, reconnaissance, threat evaluation, and even weapon control. The earliest examples of large sensor data fusion projects were designed for air defence against missiles and low-flying

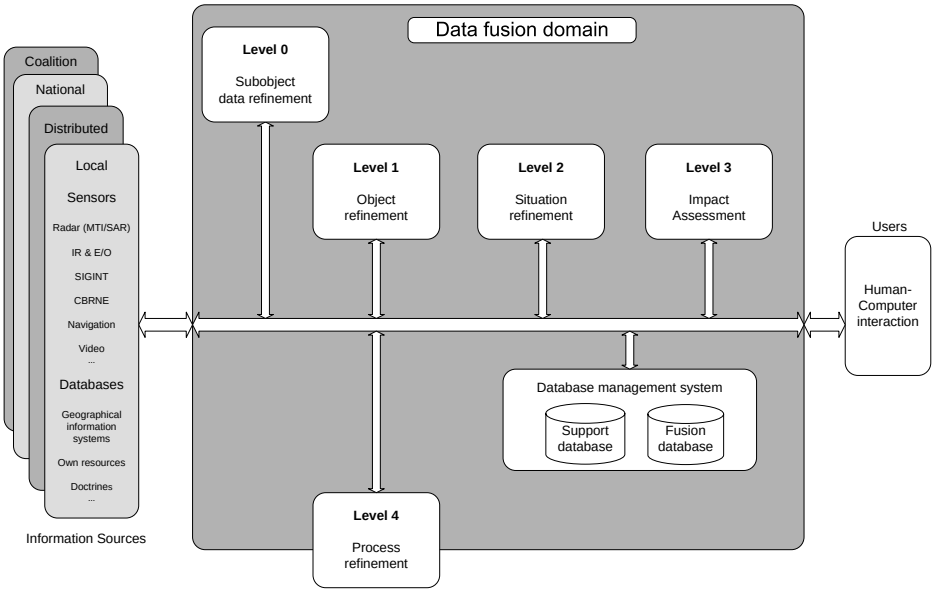


Figure 1: Overview of the JDL-Model of Sensor Data and Information Fusion [1, Chapter 3], which provides a structured and integrated view on the complete functional chain from distributed sensors, data bases, and human reports to the users and their options to act including various feed-back loops at different levels.

bombers and influenced the development of civilian air traffic control systems. The development of modern sensor data fusion technology and the underlying branch of applied science was stimulated by the advent of sufficiently powerful and compact computers and high frequency devices, programmable digital signal processors, and last but not least by the "Stratecic Defence Initiative (SDI)" announced by US President RONALD REAGAN on March 23, 1983.

After a certain level of maturity has been reached, the Joint Directors of Laboratories (JDL), an advisory board to the US Department of Defense, coined the technical term "Sensor Data and Information Fusion" in George Orwell's very year 1984 and undertook the first attempt of a scientific systematization of the new technology and the research areas related to it [1, Chapter 2, p. 24]. To the present day, the scientific fusion community speaks of the "JDL Model of Information Fusion" and its subsequent generalizations and adaptations [1, Chapter 3], [2]. The JDL model provides a structured and integrated view on the complete functional chain from distributed sensors, data bases, and human reports to the users and their options to act including various feed-back loops at different levels (Figure 1). It seems to be valid even in the upcoming large fields of civilian applications of sensor data fusion and computer security [3]. Obviously, the fundamental concepts of sensor data fusion have been developed long before their full technical feasibility and robust realizability in practical applications.

1.3 General Technological Prerequisites

The modern development of sensor data fusion systems was made possible by substantial progress in the following areas over the recent decades:

1. Advanced and robust *sensor systems*, technical equivalents of sense organs with high sensitivity or coverage are made available that may open dimensions of perception usually unaccessible to most living creatures.

2. *Communication links* with sufficient bandwidths, small latencies, stable connectivity, and robustness against interference are the backbones of spatially distributed networks of homogeneous or heterogeneous sensors.

3. Mature *navigation systems* are prerequisites of (semi-)autonomously operating sensor platforms and common frames of reference for the sensor data based on precise space-time registration including mutual alignment.

4. *Information technology* provides not only sufficient processing power for dealing with large data streams, but also efficient data base technology and fast algorithmic realizations of data exploitation methods.

5. *Technical interoperability*, the ability of two or more sub-systems or components to exchange and to information, is inevitable to build distributed "systems of systems" for sensor exploration and data exploitation [4].

6. Advanced and ergonomically efficient *human-machine interaction (HMI* tools are an integral part of man-machine-systems presenting the results of sensor data fusion systems to the users in an appropriate way [5].

The technological potential enabled by all these capabilities is much enhanced by integrating then in an overlay sensor data fusion system.

1.4 Relation to Information Systems

According to this technological infrastructure, human decision makers on all levels of hierarchy, as well as automated decision making systems, have access to vast amounts of data. In order to optimize use of this high degree of data availability in various decision tasks, however, the data continuously streaming in must not overwhelm the human beings, decision making machines, or actuators involved. On the contrary, the data must be fused in such a way that at the right instant of time the right piece of high-quality information relevant to a given situation is transmitted to the right user or component and appropriately presented. Only if this is the case, can the data streams support goal-oriented decisions and coordinated action planing in practical situations and on all levels of decision hierarchy.

In civilian applications, management information or data warehouse systems are designed in order to handle large information streams. Their equivalents in the defence and security domain are called C^4ISTAR Systems [4]. This acronym denotes computer-assisted functions for C^4 (Command, Control, Communications, Computers), I (Intelligence), and STAR (Surveillance, Target Acquisition and Reconnaissance) in order to enable the coordination of defence-related operations. While management information or data warehouse systems are primarily used to obtain competitive advantages in economic environments, C^4ISTAR systems aim at information dominance over potential opponents. The observation that more or less the same terminology is used in both areas for characterizing the struggle to avoid harm or successfully reach goals, is an indication of far-reaching fundamental commonalities of decision processes in defence command & control as well as in product development and planing, in spite of different accentuations in particular aspects.

A basic component of C^4ISTAR information systems, modular and flexibly designed as "systems of systems", is the combination of sensor systems and data bases with appropriate sensor data and information fusion sub-systems. The objective at this level is the production of timely, consistent and, above all, sufficiently complete and detailed "situation pictures", which electronically represent a complex and dynamically evolving overall scenario in the air, on the ground, at sea, or in an urban environment. The concrete operational requirements and restrictions in a given application define the particular information sources to be considered and data fusion techniques to be used.

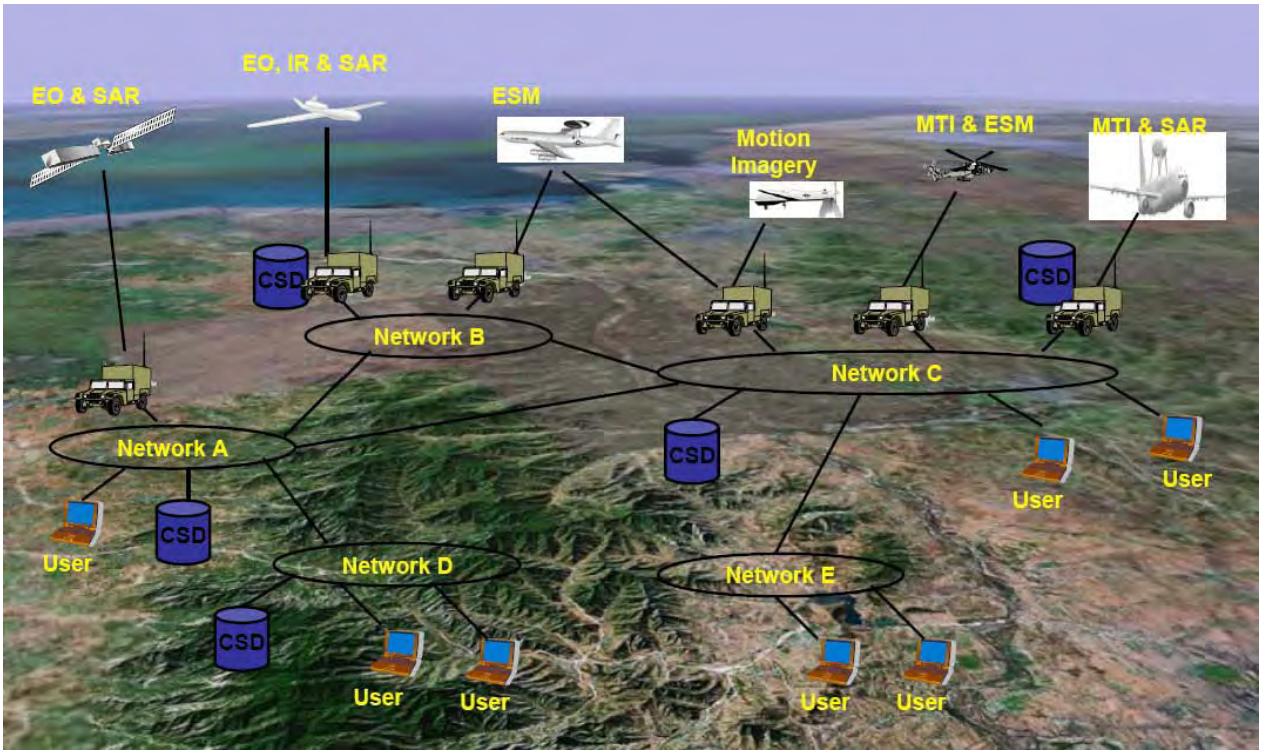


Figure 2: MAJIIC system architecture emphasizing the deployed sensors, databases, and distributed sensor data fusion systems (Interoperable ISR Exploitation Stations).

2 Discussion of a Characteristic Example

A particularly mature example of an information system, where advanced sensor data fusion technology is among its central pillars, is given by a distributed, coalition-wide C^4ISTAR system of systems for wide-area ground surveillance. It mirrors many of the aspects previously addressed and has been carried out within the framework of a multinational technology program called MAJIIC (Multi-Sensor Aerospace-Ground Joint ISR Interoperability Coalition) [4, Chapter 20]. By collaboratively using interoperable sensor and data exploitation systems in coalition operations, MAJIIC has been designed to improve situational awareness of military commanders over the various levels of the decision making hierarchy.

Based on appropriate concepts of deployment and the corresponding tactical procedures, technological tools for Collection, Coordination and Intelligence Requirements Management (CCIRM) are initiated by individual sensor service requests of deployed action forces. The CCIRM tools produce mission plans according to superordinate priorities, task sensor systems with appropriate data acquisition missions, initiate data exploitation and fusion of the produced sensor data streams in order to obtain high-quality reconnaissance information, and, last but not least, guarantee the feedback of the right information to the requesting forces at the right instant of time.

Under the constraint of leaving existing C^4ISTAR system components of the nations participating in MAJIIC unchanged as far as possible, the following aspects are addressed with particular emphasis:

1. The integration of advanced sensor technology for airborne and ground-based wide-area surveillance is mainly based on Ground Moving Target Indicator Radar (GMTI), Synthetic Aperture Radar (SAR), electro-optical and infrared sensors (E/O, IR) producing freeze and motion imagery, Electronic Support Measures (ESM), and artillery localization sensors (radar- or acoustics-based).

2. Another basic issue is the identification and implementation of common standards for distributing

sensor data from heterogeneous sources including appropriate data and meta-data formats, agreements on system architectures as well as the design and implementation of advanced information security concepts.

3. In addition to sensor data fusion technology itself, tools and procedures have been developed and are continuously enhanced for co-registration of heterogeneous sensors, cross-cueing between the individual sensors of a surveillance system, the sensors of different systems, and between sensors and actuators, as well as for exploitation product management, representation of the "Coalition Ground Picture", for coordinated mission planning, tasking, management, and monitoring of the MAJIIC subsystems.

4. MAJIIC-specific communications have been designed to be independent of network-types and communication bandwidths, making it adaptable to varying requirements. Commercially available and standardized internet- and crypto-technology has been used in both the network design and the implementation of interfaces and operational features. Important functionalities are provided by collaboration tools enabling ad-hoc communication between operators and exchange of structured information.

5. The central information distribution nodes of MAJIIC C^4ISTAR system of systems are so-called Coalition Shared Data servers (CSD) making use of modern database technology. Advanced Data Mining and Data Retrieval tools are part of all MAJIIC data exploitation and fusion systems.

6. From an operational point of view, a continuous interaction between Concept Development and Experimentation by planning, running, and analyzing simulated and live C^4ISTAR experiments is an essential part of the MAJIIC program, fostering the transfer of MAJIIC capabilities into national and coalition systems.

Figure 2 provides an overview of the MAJIIC system architecture and the deployed sensor systems.

2.1 From Imperfect Data to Situation Pictures

Sensor data fusion typically provides answers to questions related to objects of interest such as: Do object exist at all and how many of them are moving in the sensors' fields of view? Where are they geolocated at what time? Where will they be in the future with what probability? How can their overall behavior be characterized? Are anomalies or hints to their possible intentions recognizable? What can be inferred about the classes the objects belong to or even their identities? Are there clues for characteristic interrelations between individual objects? In which regions do they have their origin? What can be said about their possible destinations? Are there observable over-all object flows? Where are sources or sinks of traffic? and many other questions.

The answers to those questions are the constitutive elements, from which near real-time situation pictures can be produced that electronically represent a complex and dynamically evolving overall scenario in the air, on the ground, at sea, under water, as well as in out- or in-door urban environments, and even more abstract spaces. According to the previous discussion, these "situation elements" must be gained from the currently received sensor data streams while taking into account all the available context knowledge and pre-history. Since situation pictures are fundamental to any type of computer-aided decision support, the requirements of a given application define which particular information sources are to be fused.

The sensor data to be fused are usually inaccurate, incomplete, or ambiguous. Closely-spaced moving objects are often totally or partially irresolvable. The measured object parameters may be false or corrupted by hostile measures. The context information is in many cases hard to formalize and even contradictory in certain aspects. These deficiencies of the information to be fused are unavoidable in any real-world application. Therefore, the extraction of 'information elements' for situation pictures is by no means trivial and requires a sophisticated mathematical methodology for dealing with imperfect information. Besides a

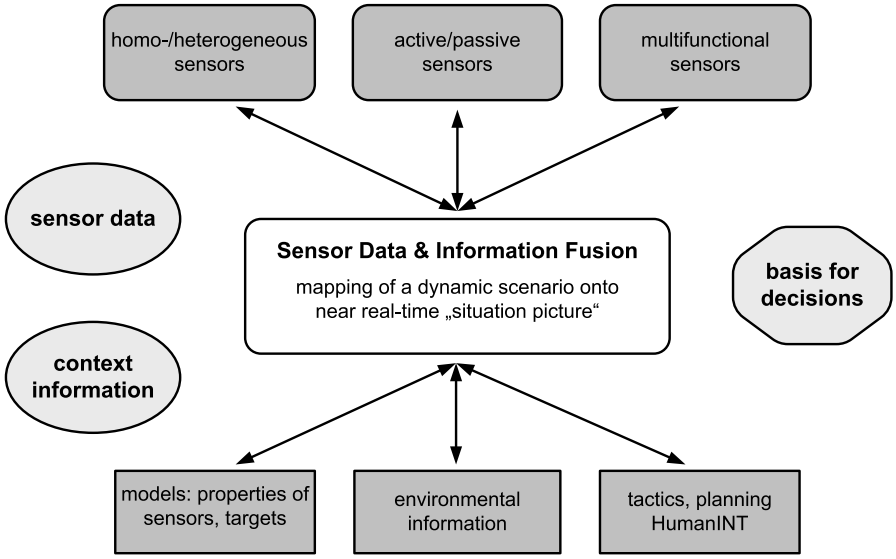


Figure 3: Sensor data and information fusion for situation pictures: overview of characteristic aspects and their mutual interrelation.

precise requirement analysis, this is one of the major scientific features that characterizes and shapes sensor data fusion as branch of applied science.

2.2 Aspects in Situation Picture Production

Figure 3 provides an overview of different aspects within this context and their mutual interrelation, which should be emphasized::

1. The underlying sensor systems can be located in different ways (collocated, distributed, mobile) producing measurements of the same or of different type. A multisensor system potentially increases the coverage or data rate of the total system and may help to resolve ambiguities.

2. Even by fusing homogeneous sensors, information can be obtained that is unaccessible to each sensor individually, such as in stereoscopic vision, where range information is provided by fusing two camera images taken from different viewpoints.

3. Fusion of heterogeneous sensor data is of particular importance, such as the combination of kinematic measurements with measured attributes providing information on the classes to which objects belongs to. Examples for measured attributes are Signal Intelligence (SIGINT), Jet Engine Modulation (JEM), radial or lateral object extension, chemical signatures etc.

4. Especially for defense and security applications, the distinction between active and passive sensing is important as passive sensors enable covert surveillance, which does not reveal itself by actively emitting radiation.

5. Multi-functional sensor systems, such as phased-array radar, offer additional operational modes, thus requiring more intelligent strategies of sensor management that provide feedback to the process of information acquisition via appropriate control or correction commands. By this, the surveillance objectives can often be reached much more efficiently.

6. Context information is given, for example, by available knowledge on sensor and object properties, which is often quantitatively described by statistical models. Context knowledge is also given by environmental information on roads or topographical occlusions and provided by Geographical Information Systems (GIS). Seen from a different perspective, context information, such as road maps, can also be extracted from real-time sensor data directly.

7. Militarily relevant context knowledge (e.g. doctrines, planning data, tactics) and human observer reports (HUMINT: Human Intelligence) is also important information in the fusion process. The exploitation of context information of this kind can significantly improve the fusion system performance.

2.3 Remarks on the Methods Used

Situation elements for producing timely situation pictures are provided by integratively and spatio-temporally processing various pieces of information that in themselves often may have only limited value for understanding the situation. Essentially, logical cross-references, inherent complementarity, and redundancy are exploited. More concretely speaking, the methods used are characterized by a stochastic approach (estimating relevant state quantities) and a more heuristically defined knowledge-based approach (modeling actual human behavior when exploiting information).

Among the data exploitation products of data fusion systems, object 'tracks' are of particular importance. Tracking faces an omnipresent aspect in every real-world application insofar as it is dealing with fusion of data produced at *different instants of time*; i.e. tracking is important in all applications where particular emphasis is placed on the fact that the sensor data to be exploited have the character of a time series.

Tracks thus represent currently available knowledge on relevant, time-varying quantities characterizing the instantaneous "state" of individual targets or target groups of interest, such as aircraft, ships, submarines, vehicles, or moving persons. Quantitative measures that reliably describe the quality of this knowledge are an integral part of a track. The information obtained by 'tracking' algorithms [6, 7, 8, 9] also includes the history of the targets. If possible, a one-to-one association between the target trajectories in the sensors' field of view and the produced tracks is to be established and has to be preserved as long as possible (track continuity). The achievable track quality does not only depend on the performance of the sensors used, but also on target properties and the operational conditions within the scenario to be observed. If tracks 'match' with the underlying real situation within the bounds defined by inherent quality measures being part of them, we speak of 'track consistency".

Tracking algorithms, including Bayesian multiple hypothesis trackers as particularly well-understood examples, are iterative updating schemes for conditional probability density functions representing all available knowledge on the kinematic state of the objects to be tracked at discrete instants of time t_l. The probability densities are conditioned by both, the sensor data accumulated up to some time t_k, typically the current data acquisition time, as well as by available context information, such as on sensor characteristics, the object dynamics, the environment, topographical maps, or on certain rules governing the object behavior. Depending on the time instant t_l at which estimates for the state $\mathbf{x}_l$ are required, the related estimation process is referred to as prediction ($t_l > t_k$), filtering ($t_l = t_k$), or retrodiction ($t_l < t_k$) [10, 11].

2.4 A Schematic Sensor Data Fusion System

Figure 4 shows a generic scheme of functional building blocks within a multiple sensor tracking and data fusion system along with its relation to the underlying sensors. In the case of multi-functional sensors, there is feedback from the tracking system to the process of sensor data acquisition (sensor management). The following aspects should be emphasized:

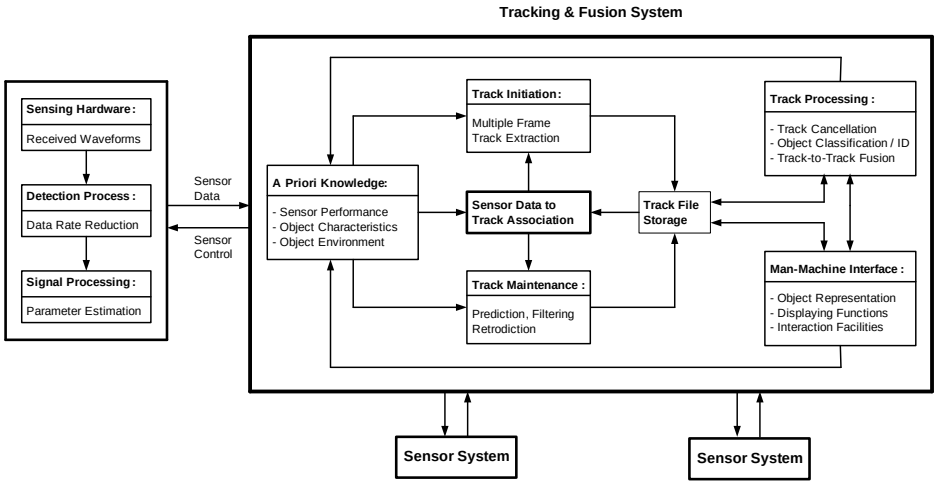


Figure 4: Generic scheme of functional building blocks within a tracking/fusion system along with its relation to the sensors (centralized configuration, type IV according to O. Drummond).

2.4.1 Sensor Systems

After passing a detection process, essentially working as a means of data rate reduction, the signal processing provides estimates of parameters characterizing the waveforms received at the sensors' front ends (e.g. radar antennas). From these estimates sensor reports are created, i.e. measured quantities possibly related to objects of interest, which are the input for the tracking and sensor data fusion system. By using multiple sensors instead of one single sensor, among other benefits, the reliability and robustness of the entire system is usually increased, since malfunctions are recognized easier and earlier and often can be compensated without risking a total system breakdown.

2.4.2 Interoperability

A prerequisite of all further processing steps, which at first sight seems to be trivial, is technical interoperability. It guarantees that all relevant sensor data are transmitted properly, in a timely way, and completely including all necessary meta-data describing the sensor performance, the platform parameters, and environmental characteristics. This type of meta data is necessary to transform the sensor data into common frames of reference, to identify identical pieces of data, and to merge similar pieces of data into one single augmented piece of information. The process of combining data from different sources and providing the user with a unified view of these data is sometimes also referred to as data integration. Often interoperability acts as a bottleneck in designing real-world data fusion systems of systems [4, Chapter 20].

2.4.3 Fusion Process

All sensor data that can be associated to existing tracks are used for track maintenance (using, e.g., prediction, filtering, and retrodiction). The remaining data are processed for initiating new tentative tracks (multiple frame track extraction). Association techniques thus play a key role in tracking/fusion applications. Context information in terms of statistical models (sensor performance, object characteristics, object environment) is a prerequisite for track maintenance and initiation. Track confirmation/termination, classification/identification, and fusion of tracks related to the same objects or object groups are part of the track management functionalities.

2.4.4 Human-Machine Interface

The scheme is completed by a human-machine interface with display and interaction functions. Context information can be updated or modified by direct human interaction or by the track processor itself, for example as a consequence of object classification or road map extraction. For an introduction to the vast literature on the related problems in human factors engineering and on practical systems solutions see [5].

2.5 On Measuring Fusion Performance

In sensor data fusion, the underlying 'real' situation is typically unknown. Only in expensive and time-consuming experiments certain aspects of a dynamically evolving situation are monitored, sometimes even with questionable accuracy. For this reason, experiments are valuable for demonstrating the "proof of concept" as well as to understand the underlying physical phenomena and operational problems, for example. They are of limited use, however, in performance evaluation and prediction. This underlines the role of comprehensive Monte-Carlo-simulations in fusion system performance evaluation.

According to the previous discussion, sensor data fusion systems try to establish one-to-one relations between objects in the sensors' fields of view and identified object tracks in the situation picture. Strictly speaking, this is only possible under ideal conditions regarding the sensor performance and the underlying target scenario. It seems thus reasonable to measure the performance of a given tracking/fusion system by its characteristic deficiencies when compared to this ideal goal. In general, two categories of deficiencies can be distinguished that are either caused by mis-match regarding the input data or by non-optimal processing and unfavorable application constraints.

Selected Performance Measures Selected performance measures or 'measures of deficiency' in the sense of the previous discussion, which have practical relevance in fusion systems design should be emphasized in the following.

1. Usually a time delay is involved until a track has been extracted from the sensor data. A corresponding performance measure is thus given by the 'extraction delay' between the first detection of a target by a sensor and a confirmed track.

2. False tracks, i.e. tracks related to unreal or unwanted targets, are unavoidable in the case of a high false return density (e.g. by clutter, jamming/deception). Corresponding 'deficiencies' are: mean number of falsely extracted targets per time and mean life time of a false track before its deletion.

3. Targets should be represented by one and the same track until leaving the field of view. Related performance measures are: mean life time of true target tracks, probability of an 'identity switch', and probability of a target not being represented by a track.

4. The track inaccuracy (given by the error covariance matrix of a state estimate, e.g.) should be as small as possible. Furthermore, the deviations between the estimated and actual target characteristics should correspond with the error covariance matrices produced (consistency). If this is not the case, 'track loss' usually occurs.

In a given application it is by no means simple to achieve a reasonable compromise between the various, competing performance measures and the user requirements. Optimization with respect to one measure may easily degrade other performance measures, finally deteriorating the entire system performance. This is especially true under more challenging conditions.

2.6 Tracking-derived Situation Elements

The primary objective of multiple sensor target tracking is to explore the underlying target kinematics such as position, velocity, or acceleration. In other words, standard target tracking applications gain information related to 'Level 1 Fusion' according to the well-established terminology of the JDL model of information fusion (see e.g. [1, Chapter 2] and the literature cited therein). Kinematic data of this type, however, are by no means the only information to be derived from target tracks. In many cases, reliable and quantitative higher level information according to the JDL terminology can be obtained. To be more concrete, wide-area air and ground surveillance is considered here as an important real-world example serving as a paradigm for other challenging tracking and fusion applications.

2.6.1 Inferences based on Retrodicted Tracks

The first type of higher JDL level information to be inferred from tracking data is based on a closer analysis of the histories of the kinematic object states provided by retrodiction techniques. The statements derived typically refer to object characteristics that are either time invariant or change with time on a much larger scale than kinematics quantities usually tend to do. This is the main reason why the gain in accuracy achievable by retrodiction techniques can be exploited.

- *Velocity History.* The analysis of precisely retrodicted velocity histories enables the distinction of objects belonging to different classes such as moving persons, boats, vehicles, vessels, helicopters, or jet aircraft. If the object speed estimated with sufficiently high accuracy has exceeded a certain threshold, certain object classes can be reliably be excluded. As an example, uncertainty whether an object is a helicopter or a wing aircraft can be resolved if in the track history a velocity vector 'Zero' exists. Depending on the context of the underlying application, classifications of this type can be essential to generate an alert report.

- *Acceleration History.* Similar considerations are valid if acceleration histories are taken into account: High normal accelerations, e.g., are a clear indication of a fighter aircraft. Moreover, one can safely conclude that a fighter aircraft observed with a normal acceleration $> 6\,g$, for example, is not carrying a certain type of weaponry (any more). In other words, conclusions on the threat level connected with the objects observed can be drawn by analyzing kinematic tracks.

- *Heading, Aspect Angle.* Precise reconstructions of the targets' heading vectors are not only important input information for threat evaluation and weapon assignment in themselves, but also enable estimates of the aspect angle of an object at a given instant of time with respect to other sensors, such as those producing high range or Doppler resolution spectra. Track-derived information of this type is basic for fusing spectra distributed in time and can greatly improve object classification thus providing higher-JDL-level information.

- *Rare Event Detection.* Analysis of JDL-level-1 tracks can be the key to detecting rare or anomalous events by fusing kinematic tracks with other context information such as annotated digital road maps and general rules of behavior. A simple example in the area of continuous-time, wide-area ground surveillance can be the production of an alert message if a large freight vehicle is observed at an unusual time on a dirt road in a forest region. There are analogous examples in the maritime or air domain.

2.6.2 Inferences based on Multiple Target Tracking

A second type of higher JDL level information related to mutual object interrelations can be inferred from JDL level 1 tracking data if emphasis is placed on the results of *multiple target* tracking.

- *Common History.* Multiple target tracking methods can identify whether a set of targets belongs to the same collectively moving group, such as an aircraft formation or a vehicle convoy, whose spatial extension may be estimated and tracked. If an aircraft formation has split off after a phase of penetration, e.g., the interrelation between the individual objects is to be preserved and provides valuable higher-JDL-level information that is important, e.g., when a former group target is classified as 'hostile' since this implies that all other targets originally belonging to the same group are likely to be hostile as well.

- *Object Sources and Sinks.* The analysis of large amounts of target tracks furthermore enables the recognition of sources and sinks of moving targets. By this type of reasoning, certain areas can be identified as air fields, for example, or an area of concentration of military forces. In combination with available context information, the analysis of multiple object tracks can also be used for target classification by origin or destination. A classification as hostile or suspect directly leads to an alert report.

- *Split-off Events.* By exploiting multiple target tracking techniques, certain split-off events can be identified as launches of air-to-air or air-to-surface missiles. The recognition of such an event from JDL-level-1 tracking information not only has implications on classifying the original target as a fighter aircraft, but can also establish a certain type of 'book-keeping', such as counting the number of missile launches. This enables estimates of the residual combat strength of the object, which has direct implications on countermeasures, e.g.

- *Stopping Events.* In the case of MTI radar (Moving Target Indicator), Doppler blindness can be used to detect the event 'A target under track has stopped.', provided this phenomenon is described by appropriate sensor models. If there is previous evidence for a missile launcher, e.g., missing data due to Doppler blindness may indicate preparation for launch with implications on potential countermeasures. In combination with other tracks, a stopping event may also establish new object interrelations, for example, when a target is waiting for another and then moving with it.

2.7 Selected Issues in Anomaly Detection

Anomaly detection can be regarded as a process of information fusion that combines incomplete and imperfect pieces of mutually complementary sensor data and context information in such a way that the attention of human decision makers or decision making systems is focused on particular events that are "irregular" or may cause harm and thus require special actions, such as exploiting more specialized sensors or initiating appropriate activities by military or security personnel [12]. Fusion-based anomaly detection thus improves situational awareness. What is actually meant by "regular" or "irregular" events is higher-level information itself that depends on the context of the underlying application. Here, it is either assumed to be a priori known or to be learned from statistical long-time analysis of typical situations.

In complex surveillance applications, we can often take advantage of context information on the sensing environment insofar as it is the stationary or slowly changing "stage" where a dynamic scenario evolves. Typical examples of such environmental information are digital road or sea-/air-lane maps and related information, which can essentially be regarded as spatial motion constraints (see Figure 5 as an illustration). In principle, this information is available by Geographical Information Systems (GIS). Another category of context information is provided by visibility models and littoral or weather maps indicating regions, where a high clutter background is to be taken into account, for example. Moreover, rather detailed planning information is often available. This category of information is not only important in mission planning or in the deployment and management of sensor systems, but can be used to decide whether an object is moving on a lane or leaving it, for example. In addition, ground-, sea- or air-lane information information can be used to improve the track accuracy of lane-moving vehicles and enhance track continuity.

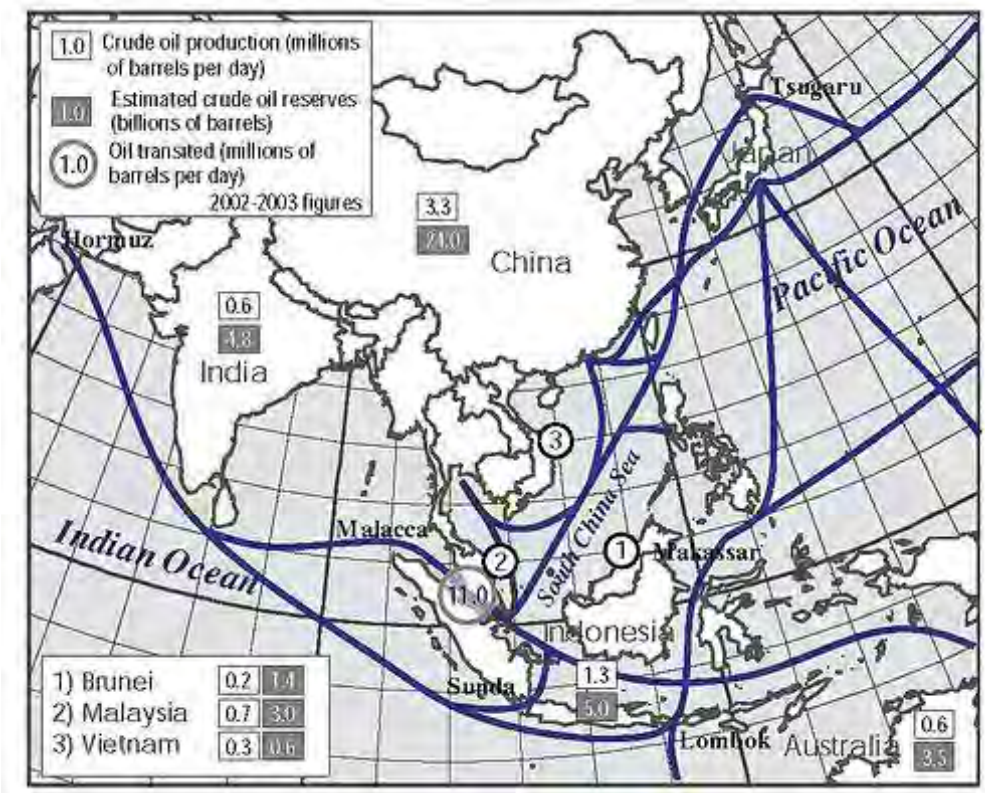


Figure 5: Illustration of sea lanes and strategic passages in Pacific Asia.

2.7.1 Integration of Planning Information

In certain applications, rather detailed planning information is available, which provides valuable context knowledge on the temporal evolution of the objects involved and can in principle be incorporated into the tracking formalism. Planning information is often approximately described by space-time waypoints that have to be passed by the individual objects during a preplanned operation, i.e. by a set of position vectors to be reached at given instants of time and possibly via particular routes (roads, lanes) between the waypoints. In addition, we assume that the acceptable tolerances related to the arrival of the objects at the waypoints are characterized by known error covariance matrices, possibly individually chosen for each waypoint and object, and that the association between the waypoínts and the objects is predefined.

The impact of waypoints on the trajectory to be estimated from future sensor data (under the assumption that the plan is actually kept) can simply be obtained by processing the waypoints as additional artificial 'measurements' via the standard Bayesian tracking paradigm, where the tolerance covariance matrices are taken into account as the corresponding 'measurement error covariances'. If this is done, the processing of sensor measurements with a younger time stamp are to be treated as "out-of sequence" measurements with respect to the artificial waypoint measurements processed earlier. According to these considerations, planning information can well improve both track accuracy and continuity as well as facilitate the sensor-data-to-track association problems involved, provided the plan is actually kept.

2.7.2 Detecting Regularity Pattern Violation

A practically important class of anomalies results from a violation of regularity patterns such as those previously discussed (motion on ground-, sea-, or air-lanes or following preplanned waypoints and routes). An anomaly detector thus has to decide between two alternatives:

- The observed objects obey an underlying pattern.

- The pattern is not obeyed (e.g. off-lane, unplanned).

Decisions of this type are characterized by decision errors of first and second. In most cases, it is desirable to make the decisions between both alternatives for given decision errors to be accepted. A "sequential likelihood ratio" test fulfills this requirement and has enormous practical importance. As soon as the test decided that the pattern is obeyed, the calculation of the likelihood ratio can be restarted since it is more or less a by-product of track maintenance. The output of subsequent sequential ratio tests can serve to re-confirm "normality" or to detect a violation of the pattern at last. The most important theoretical result on sequential likelihood ratio tests is the fact that the test has a *minimum decision length on average* given predefined statistical decision errors of first and second kind.

2.7.3 Tracking-derived Regularity Patterns

We have discussed moving targets that obey certain space-time constraints that are a priori known (roads/lanes, planned waypoints). A violation of these constraints was quite naturally interpreted as an anomaly. Seen from a different perspective, however, moving targets that are assumed to obey a priori *unknown* space-time constraints and to be observed by wide-area sensors, such as vehicles on an unknown road network, produce large data streams that can also be used for extracting the underlying space-time constraint, e.g. a road map. After a suitable post-processing, the produced tracks of motion-constrained targets simply define the corresponding constraints and can thus be extracted from tracking-based results. Extracted road-maps can be highly up-to-date and precise. A discussion where such ideas are used in wide-area maritime surveillance using AIS data can be found in [13] (AIS: Automatic Identification System).

3 Integration of Advanced Sensor Properties

Advanced signal processing techniques exploit even sophisticated physical phenomena of objects of interest and are fundamental to modern sensor system design. In particular, they have a direct impact on the quantitative and qualitative properties of the sensor data produced and to be fused. This makes a more subtle modeling of the statistical characteristics of the sensor output inevitable. Via constructing appropriate sensor models based on a deeper insight into the physical and technical sensor design principles, the performance of tracking and sensor data fusion systems can be significantly improved.

This section is focused on selected physical and technical properties of sensor systems that are used in real-world ISR applications (Intelligence, Surveillance, and Reconnaissance), such as those discussed in [4, Chapter 20]. The analysis of characteristic examples shows that context information on particular performance features of the sensor systems involved is useful, in some cases even inevitable, to fulfill an overall ISR task. The Bayesian methodology discussed in Part I is wide and flexible enough to integrate more sophisticated, appropriately designed, but still mathematically tractable likelihood functions into the process of Bayesian Knowledge Propagation. The discussed examples cover finite sensor resolution and main-lobe jamming.

The possibility to exploit even *negative sensor evidence* is a consequence that is directly connected with the use of more advanced sensor models. This notion covers the conclusions to be drawn from expected, but actually missing sensor measurements for improving the state estimates of objects under track. Even a failed attempt to detect an object of interest is a useful sensor output that is interpretable only if a consistent sensor modeling is available.

3.1 Finite Sensor Resolution

Air surveillance in a dense object / dense clutter environment is a difficult task that requires refined data association and tracking techniques. In this context, tracking for maneuvering groups of objects that join, operate closely-spaced for a while, and split off again is confronted with mainly three problems:

1. *Sensor Resolution:* Due to the limited resolution of every radar sensor, closely-spaced targets will continuously transition from being resolved to irrresolved and back again. The importance of resolution phenomena has been addressed in [14].

2. *Data Association:* Ambiguous data-to-object associations due to overlapping expectation gates are an inherent problem for formations, which is made even more difficult by high false return densities and missed detections.

3. *Maneuvers:* Often distinct maneuvering phases can be identified, as even agile objects will not always make use of their maneuvering capability. Nevertheless, abrupt turns may occur, e.g. if a formation dissolves into well-separated objects.

These problems require the use of multiple hypothesis, multiple model tracking methods as discussed in Part I. The multiple hypothesis character mirrors the uncertain origin of the data, while the multiple models refer to the different maneuvering phases. In [15], we proposed a model providing a qualitative description of resolution phenomena in terms of the *resolution probability*, by which potentially irresolved measurements can be handled within the Bayesian framework. The data association problem is covered by a likelihood function $p(Z_k, m_k|\mathbf{x}_k)$ that statistically describes what a set of m_k observations $Z_k = \{\mathbf{z}_k^j\}_{j=1}^{m_k}$ can say about the joint state $\mathbf{x}_k$ of the objects to be tracked. Due to the Total Probability Theorem, it can be written as a sum over all possible, mutually exclusive, and exhaustive data interpretations j_k:

$$p(Z_k, m_k|\mathbf{x}_k) = \sum_{j_k} p(Z_k, m_k, j_k|\mathbf{x}_k) \tag{1}$$

$$= \sum_{j_k} p(Z_k|m_k, j_k, \mathbf{x}_k)\, p(m_k|j_k, \mathbf{x}_k)\, p(j_k|\mathbf{x}_k). \tag{2}$$

Generally, the formulation of such likelihood functions is by no means a trivial task. In many practical cases, however, a given multiple-object tracking problem can be decomposed into independent sub-problems of reduced complexity. The example below is practically important but can still be handled more or less rigorously.

3.1.1 A Radar Resolution Model

For the sake of conciseness, we confine the discussion to non-imaging radar sensors. With some modifications, the results can also be transfered to infrared or electro-optical sensors, for example. Let us consider a medium range radar producing range and azimuth measurements of an object formation consisting of two targets. For physical reasons the resolution in range, azimuth, and range-rate will be independent from each other. In particular, range and cross-range resolution differ significantly in many radar applications. Therefore, the resolution performance of the sensor is expected to depend strongly on the current sensor-to-group geometry and the relative orientation of the targets within the group. The sensor's resolution capability is also determined by the particular signal processing techniques used and the random target fluctuations. As a complete description is rather complicated, we have to look for a simplified, but qualitatively correct and mathematically tractable model. In any case, the radar resolution capability in range and azimuth is limited by the corresponding band- and beam-width. These radar-specific parameters must explicitly enter into any processing of potentially irresolved plots. The typical size of resolution cells in a medium distance is about 50 m (range) and 500 m (cross range). As in target formations the mutual distance may well be 50 - 500 m or even less, the limited sensor resolution is a real problem in object tracking.

Centroid Measurements Under the hypothesis $j_k = E_k^{ii}$ assuming that the radar plot $\mathbf{z}_k^i$ is an irresolved measurement belonging to two targets with a joint vector $\mathbf{x}_k = (\mathbf{x}_k^{1\top}, \mathbf{x}_k^{2\top})^\top$, the conditional likelihood is

given by:

$$p(\mathbf{z}_k^i|\mathbf{x}_k) = \mathcal{N}\big(\mathbf{z}_k^i;\, \mathbf{H}_k^g\mathbf{x}_k,\, \mathbf{R}_k^g\big), \tag{3}$$

where the measurement matrix $\mathbf{H}_k^g$ describes a centroid measurement of the group center, characterized by a corresponding measurement error covariance matrix $\mathbf{R}_k^g$:

$$\mathbf{H}_k^g\mathbf{x}_k = \tfrac{1}{2}\mathbf{H}_k(\mathbf{x}_k^1 + \mathbf{x}_k^2), \tag{4}$$

where $(r_k, \varphi_k)^\top = \mathbf{H}_k\mathbf{x}_k^i$, $i = 1, 2$, is the measurement of the underlying tracking problem, where resolution phenomena are irrelevant.

Resolution Probability Resolution phenomena will be observed if the range and angular distances between the objects are small compared with α_r, α_φ: $\Delta r_k/\alpha_r < 1$ and $\Delta\varphi_k/\alpha_\varphi < 1$. The objects within the group are resolvable if $\Delta r_k/\alpha_r \gg 1$ or $\Delta\varphi_k/\alpha_\varphi \gg 1$. Furthermore, we expect a narrow transient region. A more quantitative description is provided by introducing a resolution probability $P_r = P_r(\Delta r, \Delta\varphi)$ depending on the sensor-to-group geometry. It can be expressed by a corresponding probability of being irresolvable $P_r = 1 - P_u(\Delta r_k, \Delta\varphi_k)$. Let us describe P_u by a Gaussian-type function of the relative range and angular distances [15]:

$$P_u(\Delta r_k, \Delta\varphi_k) = \exp\left[-\log 2(\tfrac{\Delta r_k}{\alpha_r})^2\right]\, \exp\left[-\log 2(\tfrac{\Delta\varphi_k}{\alpha_\varphi})^2\right]. \tag{5}$$

Obviously, this simple model for describing resolution phenomena reflects the previous, more qualitative discussion. We in particular observe that P_u is reduced by a factor of 2 if Δr_k is increased from zero to α_r. Due to the Gaussian character of its dependency on the state vector $\mathbf{x}_k$ the probability P_u can formally be written in terms of a normal density:

$$P_u = \exp\left[-\log 2\,\left(\mathbf{H}_k(\mathbf{x}_k^1 - \mathbf{x}_k^2)\right)^\top \mathbf{A}^{-1}(\mathbf{H}_k\mathbf{x}_k^1 - \mathbf{H}_k\mathbf{x}_k{}^2)\right] \tag{6}$$

$$= \exp\left[-\log 2\,(\mathbf{H}_k^u\mathbf{x}_k)^\top \mathbf{A}^{-1}\mathbf{H}_k^u\mathbf{x}_k\right]. \tag{7}$$

Here the *resolution matrix* $\mathbf{A}$ is defined by $\mathbf{A} = \mathbf{diag}(\alpha_r^2, \alpha_\varphi^2)$, while the quantity $\mathbf{H}_k^u\mathbf{x}_k = \mathbf{H}_k(\mathbf{x}_k^1 - \mathbf{x}_k^2)$ can be interpreted a measurement matrix for distance measurements. Up to a constant factor the resolution probability probability $P_u(\mathbf{x}_k)$ might formally be interpreted as the fictitious likelihood function of a measurement 0 of the distance $\mathbf{H}_k(\mathbf{x}_k^1 - \mathbf{x}_k^2)$ between the objects with a corresponding *fictitious* measurement error covariance matrix $\mathbf{R}_u$ defined by the resolution parameters α_r, α_φ.

$$P_u(\mathbf{x}_k) = |2\pi\mathbf{R}_u|^{-1/2}\, \mathcal{N}\big(\mathbf{O};\, \mathbf{H}_u\mathbf{x}_k,\, \mathbf{R}_k^u\big). \tag{8}$$

with $\mathbf{R}_k^u = \frac{\mathbf{A}}{2\log 2} = \frac{1}{2\log 2}\mathrm{diag}[\alpha_r^2, \alpha_\varphi^2]$. According to a first order Taylor expansion, the resolution matrix describing the resolution cells in Cartesian coordinates proves to be time dependent and results from the matrix $\mathbf{A}$ by applying dilatation and a rotation. In the same way as the Cartesian measurement error ellipses, the Cartesian "resolution ellipses" depend on the target range. Suppose we have α_r = 100 m and α_φ = 1°. We then expect that the resolution in a distance of 50 km is about 100 m (range) and 900 m (cross range). Since for military targets in a formation their mutual distance may well be 200 - 500 m or even less, resolution is a real target tracking problem.

Impact of Sensor-to-Object Geometry We expect that the resolution performance of the sensor is highly dependent on the current sensor-to-group geometry and the relative orientation of the targets within the group. As an example, let us consider the simplified situation in Figure 6. A formation with two targets is

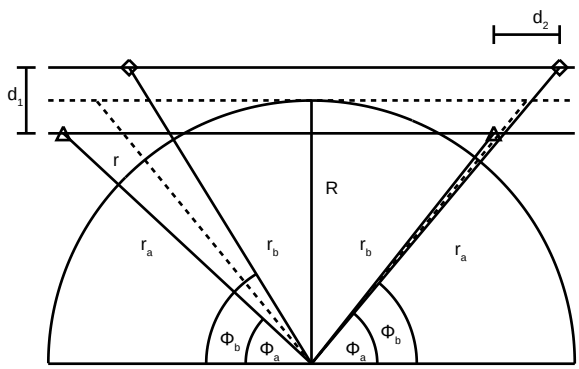


Figure 6: Radar resolution phenomena: simulated object group passing a radar sensor (left: limited by azimuth resolution, right: limited by range resolution).

passing a radar. We here consider an echelon formation. R is the minimum distance of the group center from the radar.

Figure 7 shows the resulting probability $P_u(r; R)$ parameterized by R = 0, 10, 30, 60 km as a function of the distance r between the formation center and the radar. The solid lines refer to a formation approaching the radar ($\dot{r} < 0$), the dashed lines refer to $\dot{r} > 0$. For $R \neq 0$, both flight phases differ substantially. Near R, the probability P_u varies strongly (0.85 → 0.15). For a radial flight (R = 0), we observe no asymmetry and P_u is constant over a wide range ($r \gg r_c$).

3.1.2 Resolution-specific Likelihood

For a cluster of two closely-spaced objects moving in a cluttered environment five different classes of data interpretations exist [15]:

1. $E_k^{ii}, i = 1, \ldots, m_k$: Both objects have not been resolved but detected as a group with probability P_D^u, $\mathbf{z}_k^i \in Z_k$ representing the centroid measurement; all remaining returns are false (m_k data interpretations):

$$p(Z_k|m_k, E_k^{ii}, \mathbf{x}_k) = \frac{\mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_k^g \mathbf{x}_k, \mathbf{R}_k^g)}{|\text{FoV}|^{m_k - 1}} \tag{9}$$

$$p(m_k|E_k^{ii}, \mathbf{x}_k) = p_F(n_k - 1) \tag{10}$$

$$P(E_k^{ii}|\mathbf{x}_k) = \tfrac{1}{m_k}\, P_u(\mathbf{x}_k)\, P_D^u. \tag{11}$$

With P_u as represented in Equation 8, $p(Z_k, m_k, E_k^{ii}|\mathbf{x}_k)$ is up to a constant factor given by:

$$p(Z_k, m_k, E_k^{ii}|\mathbf{x}_k) \propto \mathcal{N}\Big(\begin{pmatrix} \mathbf{z}_k^i \\ 0 \end{pmatrix}; \begin{pmatrix} \mathbf{H}_k^g \\ \mathbf{H}_k^u \end{pmatrix} \mathbf{x}_k, \begin{pmatrix} \mathbf{R}_k^g & \mathbf{O} \\ \mathbf{O} & \mathbf{R}_k^u \end{pmatrix} \Big). \tag{12}$$

Hence, under the hypothesis E_k^{ii} two measurements are to be processed: the (real) plot $\mathbf{z}_k^i$ of the group center $\mathbf{H}_k^g \mathbf{x}_k = \frac{1}{2}\mathbf{H}_k(\mathbf{x}_k^1 + \mathbf{x}_k^2)$ and a (fictitious) measurement 'zero' of the distance $\mathbf{H}_k^u \mathbf{x}_k = \mathbf{H}_k(\mathbf{x}_k^1 - \mathbf{x}_k^2)$ between the objects. We can thus speak of *'negative' sensor information* [16], as the lack of a second target measurement conveys information on the target position. In the case of a resolution conflict, the relative target distance must be smaller than the resolution.

2. E_k^0: Both objects were neither resolved nor detected as a group, so all returns in Z_k are thus assumed to be false (one interpretation hypothesis):

$$p(Z_k, m_k|E_k^0, x_k) = P_u(\mathbf{x}_k)\,(1 - P_D^u)\, p_F(m_k) \tag{13}$$

$$P(E_k^0|\mathbf{x}_k) = P_u(\mathbf{x}_k)\,(1 - P_D^u). \tag{14}$$

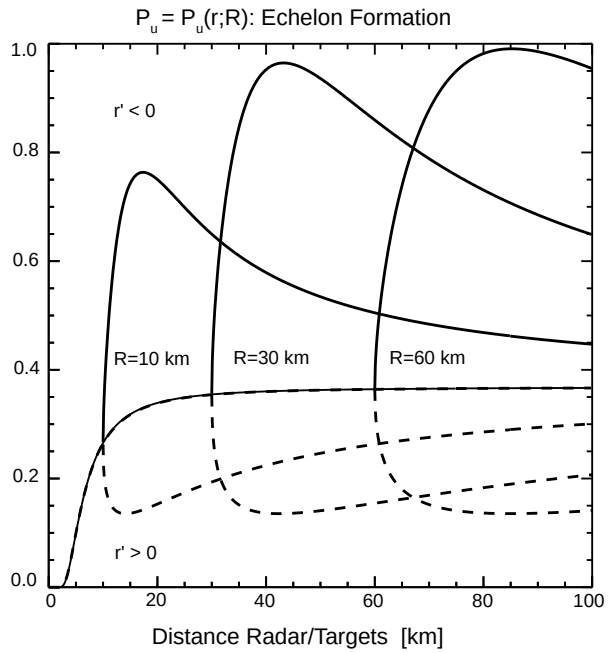


Figure 7: Effect of the underlying sensor-to-group geometry: resolution probability depending on the distance between group center and radar for R = 0, 10, 30, 60 km.

In analogy to the previous considerations, we can write up to a constant factor:

$$p(Z_k, m_k, E_k^0|\mathbf{x}_k) \propto \mathcal{N}\big(0;\, \mathbf{H}_k^u\mathbf{x},\, \mathbf{R}_k^u\big). \tag{15}$$

This means that even under the hypothesis of a missing irresolved plot, at least a fictitious distance measurement 0 is processed with a measurement error given by the sensor resolution.

3. E_k^{ij}, $i, j = 1, \ldots, m_k$, $i \neq j$: Both objects were resolved and detected, $\mathbf{z}_k^i, z_k^j \in Z_k$ are the measurements, $m_k - 2$ returns are false ($m_k(m_k - 1)$ interpretations):

$$p(Z_k|m_k, E_k^{ij}, \mathbf{x}_k) = \frac{\mathcal{N}(\mathbf{z}_k^i; \mathbf{H}_k\mathbf{x}_k^1, \mathbf{R}_k)\, \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_k\mathbf{x}_k^2, \mathbf{R}_k)}{|\mathrm{FoV}|^{m_k-2}} \tag{16}$$

$$p(m_k|E_k^{ij}, \mathbf{x}_k) = p_F(m_k - 2) \tag{17}$$

$$P(E_k^{ij}|\mathbf{x}_k) = \frac{\big(1-P_u(\mathbf{x}_k)\big)}{m_k(m_k-1)}\, P_D^2. \tag{18}$$

According to the factor $1-P_u(\mathbf{x}_k) = 1-|2\pi\mathbf{R}_u|^{\frac{1}{2}}\ \mathcal{N}\big(0;\, \mathbf{H}_k^u\mathbf{x},\, \mathbf{R}_k^u\big)$ the likelihood function becomes a mixture, in which *negative* weighting factors can occur. Nevertheless, the coefficients sum up to one; the density $p(\mathbf{x}_k|Z^k)$ is thus well-defined. This reflects the fact that in case of a resolved group the targets must have a certain minimum distance between each other which is given by the sensor resolution. Otherwise they would not have been resolvable.

4. E_k^{i0}, E_k^{0i}, $i = 1, \ldots, m_k$: Both objects were resolved but only one object was detected, $\mathbf{z}_k^i \in Z_k$ is the measurement, $m_k - 1$ returns in Z_k are false ($2m_k$ interpretations):

$$p(Z_k, m_k|E_k^{i0}, \mathbf{x}_k) = |\mathrm{FoV}|^{1-n_k}\ \mathcal{N}(\mathbf{z}_k^i;\, \mathbf{H}_k\mathbf{x}_k^1,\, \mathbf{R}_k)\, p_F(m_k - 1) \tag{19}$$

$$P(E_k^{i0}|\mathbf{x}_k) = \tfrac{1}{m_k}\, \big(1 - P_u(\mathbf{x}_k)\big)\, P_D\, (1 - P_D). \tag{20}$$

5. E_k^{00}: The objects were resolved, but not detected; all m_k plots in Z_k are false (one interpretation):

$$p(Z_k, m_k | E_k^{00}, x_k) = |\text{FoV}|^{-m_k}\, p_F(m_k) \tag{21}$$

$$P(E_k^{00} | \mathbf{x}_k) = \big(1 - P_u(\mathbf{x}_k)\big)\,(1 - P_D)^2. \tag{22}$$

Since there exist $(m_k + 1)^2 + 1$ interpretation hypotheses, the ambiguity for even small clusters of closely-spaced objects is much higher than in the case of well-separated objects ($m_k + 1$ each). This means that only small groups can be handled more or less rigorously. For larger clusters (raids of military aircraft, for instance) a collective treatment seems to be reasonable until the group splits off into smaller sub-clusters or individual objects.

Up to a factor $\frac{1}{m_k!}\,\rho_F^{m_k-2}\,|FoV|^{-m_k}\,\mathrm{e}^{-|\text{FoV}|\rho_F}$ independent of $\mathbf{x}_k$, the likelihood function of potentially irresolved sensor data in a clutter background,

$$p(Z_k, m_k | x_k) = p(Z_k, m_k, E_k^0) + \sum_{i,j=0}^{m_k} p(Z_k, E_k^{ij}, m_k | x_k), \tag{23}$$

is proportional to a sum of Gaußians and a constant:

$$\begin{aligned} p(Z_k, n_k | \mathbf{x}_k) \propto\; & \rho_F^2 (1 - P_D)^2 \big(1 - P_u(\mathbf{x}_k)\big) + \rho_F (1 - P_D^u) P_u(\mathbf{x}_k) + \\ & P_D^u \rho_F P_u(\mathbf{x}_k) \sum_{i=1}^{n_k} \mathcal{N}(\mathbf{z}_k^i;\, \mathbf{H}_k^g \mathbf{x}_k,\, \mathbf{R}_k^g) + \\ & \rho_F P_D (1 - P_D)\big(1 - P_u(\mathbf{x}_k)\big) \sum_{i=1}^{n_k} \big\{ \mathcal{N}(\mathbf{z}_k^i;\, \mathbf{H}_k \mathbf{x}_k^1,\, \mathbf{R}_k) + \mathcal{N}(\mathbf{z}_k^i;\, \mathbf{H}_k \mathbf{x}_k^2,\, \mathbf{R}_k) \big\} + \\ & P_D^2 \big(1 - P_u(\mathbf{x}_k)\big) \sum_{\substack{i,j=1 \\ i \neq j}}^{n_k} p_k^{ij}(\mathbf{x}_k)\; \mathcal{N}(\mathbf{z}_k^i;\, \mathbf{H}_k \mathbf{x}_k^1,\, \mathbf{R}_k)\; \mathcal{N}(\mathbf{z}_k^j;\, \mathbf{H}_k \mathbf{x}_k^2,\, \mathbf{R}_k). \end{aligned} \tag{24}$$

3.1.3 A Formation Tracking Example

If the spatial false return density is not too high, JPDA-type approximations [8] can be applied. According to this philosophy, the joint state mixture density $p(\mathbf{x}_k^1, \mathbf{x}_k^2 | \mathcal{Z}^k)$ resulting from the likelihood function previously discussed is approximated by a single Gaussian with the same expectation vector and covariance matrix as the mixture $p(\mathbf{x}_k^1, \mathbf{x}_k^2 | \mathcal{Z}^k)$ (moment matching [8, p. 56 ff]). Objects moving closely-spaced for some time irreversibly lose their identity: When they dissolve again, a unique track-to-target association is impossible. It is thus reasonable to deal with densities that are symmetric under permutations of the individual targets. Thus, no statistically relevant information is lost and the filter performance remains unchanged, while the mean number of hypotheses involved may be significantly reduced. Within the framework of JPDA-type approximations, this has the following effect: Before combining two components of the mixture via moment matching, we check if the components are more 'similar' to each other when the target indices are switched. If this is the case, we combine them instead. These considerations are also a useful and simple means to avoid track coalescence.

Figure 8 shows a set of data from a typical medium-range radar. The scan interval is 5 sec and the detection probability about 80%. The example clearly shows that resolution must be taken into account as soon as the targets begin to move closely-spaced. Figures 9, 10 show the estimation error ellipses for two targets (red, white) resulting from JPDA filtering. While in Figure 9 perfect sensor resolution was assumed (wrongly!), in Figure 10 the previous resolution model was used. JPDA filtering without considering resolution phenomena evidently fails after a few frames, as indicated by diverging tracking error ellipses. This has a simple explanation: without modeling the limited sensor resolution, an actually produced irresolved plot

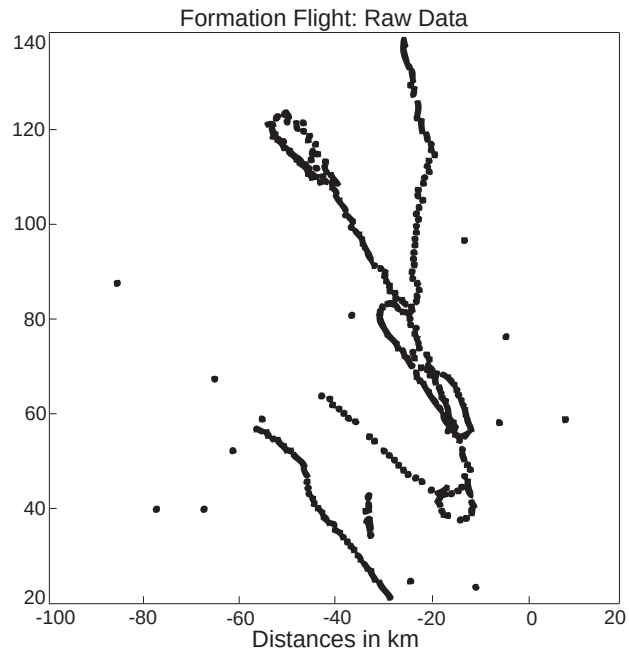


Figure 8: Partly irresolved aircraft formation: accumulated raw data of a mid-range radar.

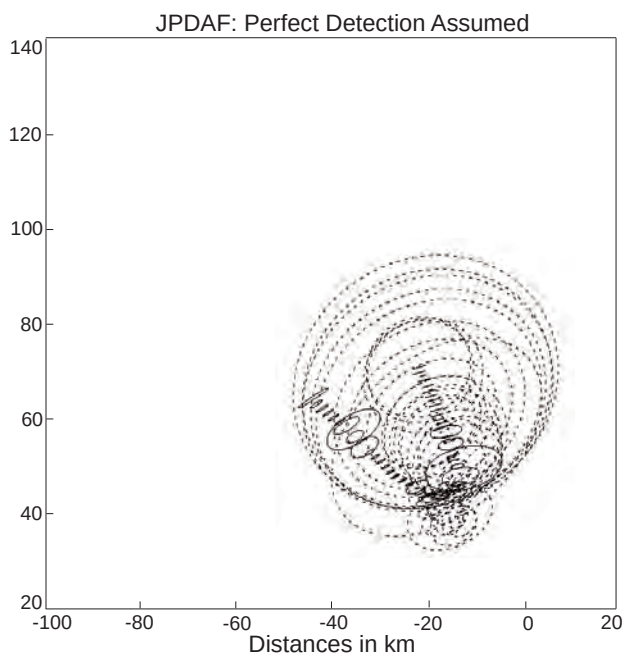


Figure 9: Tracking of an aircraft formation: filtering results (JPDA, no resolution model).

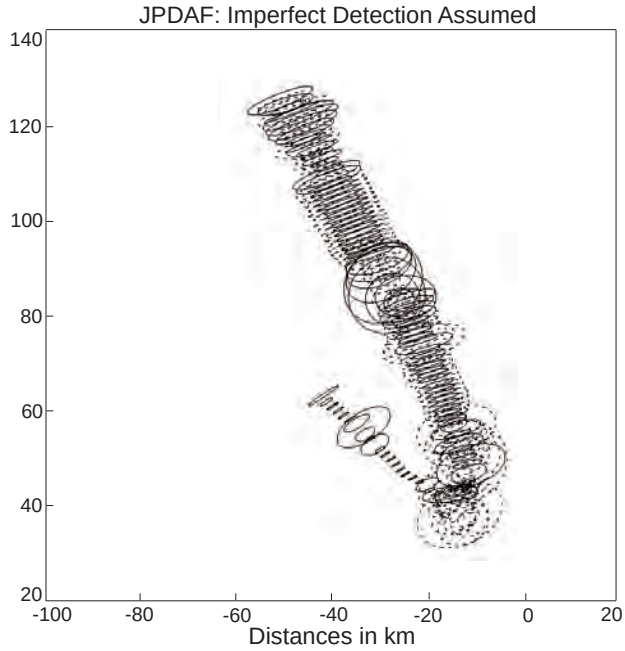


Figure 10: Tracking of an aircraft formation: filtering results (with resolution model).

can only be treated as a single target measurement along with a missed detection. In consequence, the related covariance matrices increase in size. This effect is further intensified by subsequent irresolved returns. If hypotheses related to resolution conflicts are taken into account, however, the tracking remains stable. The error ellipses in Figures 9, 10 have been enlarged to make their data-driven adaptivity more visible. The ellipses shrink, for instance, if both targets are actually resolved in a particular scan. The transient enlargement halfway during the formation flight is caused by a crossing target situation.

3.1.4 Resolution: Summary of Results

MHT filtering with explicit handling of resolution conflicts can successfully be applied to real radar data [17]. The main conclusions of extensive simulations based on exemplary scenarios and typical radar parameters are [18]:

1. For objects with overlapping expectation gates and potentially irresolved measurements, MHT filters that handle data association conflicts in combination with resolution phenomena by far outperform more conventional trackers (monohypothesis approximations or filters ignoring imperfect resolution). Much higher false return densities and significantly lower detection probabilities can be tolerated, the tracks are more accurate, the correlation gates are reduced in size, and the critical phases of joining and splitting-off are supported.

2. Provided only primary radar data are available, information on the object identity rapidly fades out while the objects move closely-spaced and produce potentially irresolved plots. After splitting off again, a unique track-to-target correlation is no longer possible. We may thus drop the notion of identity and deal with indistinguishable targets. By this, no statistically relevant information is lost, i.e. the number of hypotheses involved can significantly be reduced without affecting the track accuracy.

3. Whether an object group is resolvable or not is highly dependent on the specific sensor-to-object geometry considered and on the position of the objects relative to each other. This phenomenon is adaptively taken into account by the resolution model used. As the correct association hypotheses

can reliably be reconstructed by retrodiction techniques at the expense of some delay, the resolution model may in a retrospective view provide information on the relative position of the targets within the formation.

4. Besides the ambiguity due to irresolved or missed detections, overlapping correlation gates, and false returns, scenarios with highly maneuvering targets are also ambiguous with respect to the object evolution model assumed to be in effect. Hypotheses related to resolution conflicts fit well into the more complex framework of IMM-MHT and provide performance improvements over more simplified dynamics models.

A detailed discussion of this approach has been published in:

- W. Koch and G. van Keuk

  Multiple Hypothesis Track Maintenance with Possibly Unresolved Measurements

  IEEE Transactions on Aerospace and Electronic Systems, Vol. 33, No. 3, p.883-892, July 1997.

  An extended version with results from various related conference papers of the author has been published as a handbook chapter in: W. Koch. Target Tracking. Chapter 8 in: *Stergios Stergiopoulos (Ed.). Advanced Signal Processing: Theory and Implementation for Sonar, Radar, and Non-Invasive Medical Diagnostic Systems*. CRC Press (2001).

  Abstract
  In surveillance problems, dense clutter/dense target situations call for refined data association and tracking techniques. In addition, closely-spaced targets may exist which are not resolved. This phenomenon has to be considered explicitly in the tracking algorithm. We concentrate on two targets that temporarily move in close formation and derive a generalization of MHT methods on the basis of a simple resolution model.

  Key words: sensor resolution, Bayesian multiple target tracking, multiple hypothesis tracking, target formations

3.2 Tracking in Presence of Main-lobe Jamming

The degrees of freedom available in applications with airborne phased-array radar enable suppression of so called main-lobe jammers that try to blind the radar by transmitting specially designed radiation directly into the main beam of the radar, by using adaptive array signal processing techniques [19]. Following the spirit of the discussions in the previous sections, the current position of the resulting jammer notch as well as information on the distribution of the related monopulse measurements can be incorporated into a more sophisticated sensor performance model of air-borne phased-array radar. The proposed model does not only improve object tracking in the vicinity of a jammer notch in terms of a shorter extraction delay, improved track accuracy/continuity. It also has strong impact on strategies for adaptive sensor control.

3.2.1 Modeling the Jammer Notch

Tracking of an approaching missile under mail-lobe jamming conditions is among the most challenging data fusion tasks [20]. Advance sensor models can contribute to their efficient and robust solution. An example is the simulated situation in Figure 11, which shows the trajectories of a sensor (AESA: Active Electronically Scanned Array) on a moving platform (black), of an object to be tracked (red), and the jammer (magenta).

By using adaptive digital beamforming techniques, AESA radars of modern interceptor aircraft are able to electronically produce a sector of vanishing susceptibility in their receive beam pattern. Excepting this “blind spot”, also called jammer notch, the radar is operating more or less normally. A non-cooperative

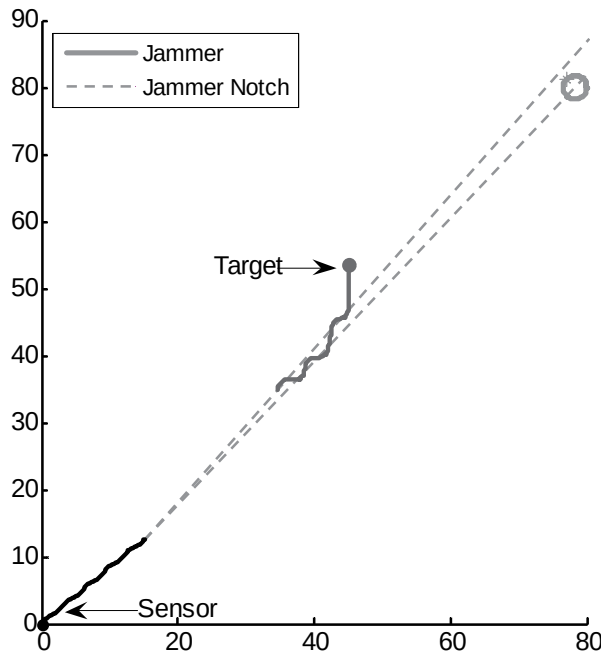


Figure 11: Moving aircraft under mail-lobe jamming conditions: approaching missile near the shadow of the jammer notch

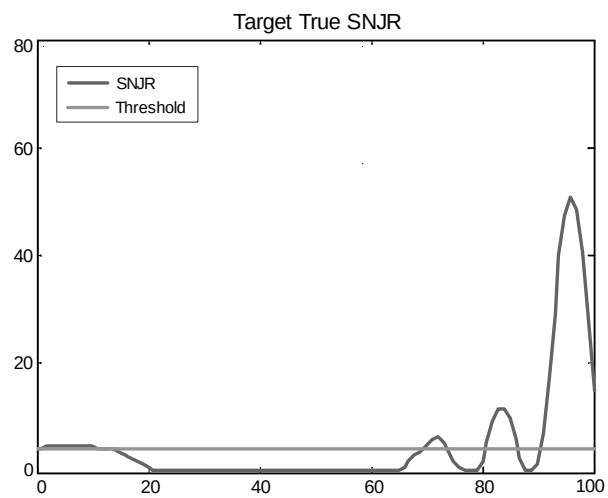


Figure 12: Temporal variation of the signal-to-noise ratio under of an approaching missile under main-lobe jamming

missile, however, is expected to approach the interceptor aircraft as long as possible in the shadow of the jammer notch. The dashed lines in Figure 11 characterize the spatial region of the blind spot depending on the current sensor-to-jammer geometry object.

The effect of the jammer is directly visible in the signal-to-noise-plus-jammer ration (SNJR) of the target, which is shown Figure 12 for the scenario discussed as a function of time. Only in the beginning can the missile be detected for a short time. Then it is masked for a long time by the radar's blind spot, until it becomes visible again in close vicinity of the sensor, where the reflected signal is very strong (Burn Through). Sophisticated signal processing provides estimates of the missile direction by using adaptive monopulse techniques [19] as well as the corresponding estimation error covariance matrix $\mathbf{R}(\mathbf{b}_k, \mathbf{j}_k)$ as an additional sensor output. $\mathbf{R}(\mathbf{b}_k, \mathbf{j}_k)$ depends on the current beam direction $\mathbf{b}_k$ of the AESA radar and the jammer direction $\mathbf{j}_k$ and describes in particular the mutual correlation of the estimated direction cosines in the vicinity of the jammer notch. It thus provides valuable context information on the sensor performance.

The sensor model is based on an expression for the signal-to-noise+jammer ratio (SNJR) after completing the signal processing chain. The following simple formula mirrors all relevant phenomena observed:

$$\mathrm{SNJR}(\mathbf{d}_k, r_k; \mathbf{b}_k, \mathbf{j}_k) \quad = \quad \mathrm{SNR}_0 \ \left(\tfrac{r_k}{r_0}\right)^{-4} \ D(\mathbf{d}_k) \ \mathrm{e}^{-\log 2|\mathbf{d}_k - \mathbf{b}_k|^2/b^2} \ \left(1 \ - \ \mathrm{e}^{-\log 2|\mathbf{d}_k - \mathbf{j}_k|^2/j^2}\right). \quad (25)$$

The vectors $\mathbf{b}_k$ and $\mathbf{j}_k$ denote the angular position of the current beam and the jammer, respectively (assumed to be known). b is a measure of the beam width, while j indicates the width of the jammer notch produced by adaptive nulling, and r_0 is the radar's instrumented range. $\mathbf{d}_k$ is the object's direction vector and r_k its range from the sensor. $D(\mathbf{d}_d)$ reflects the antenna's directivity pattern. In the case of Swerling I fluctuations of the objects' radar cross section and for a simple detection model, the detection probability is a function of $\mathbf{d}_k$, r_k, $\mathbf{b}_k$, and $\mathbf{j}_k$:

$$P_D(\mathbf{d}_k, r_k; \mathbf{b}_k, \mathbf{j}_k) = P_F^{\frac{1}{1+\mathrm{SNJR}(\mathbf{x}_k; \mathbf{b}_k, \mathbf{j}_k)}}. \quad (26)$$

P_D can be approximated by using Gaußians linearly depending on the object state. Essentially, we enter this expression of the detection probability into the likelihood function, yielding a Gaußian sum type expression for it.

3.2.2 Tracking Filters Alternatives

According to the previous discussion, the signal-to-noise-plus-jammer is essential in the modeling of the detection probability and thus enters into the likelihood function ratio. After some approximations, the likelihood function can be represented by a Gaußian mixture, finally leading to a version of the Gaußian sum filter. Since the number of mixture components grows in each update step, adaptive approximation schemes must be applied. By using Monte-Carlo-simulations five competing approaches have been evaluated and compared with each other:

1. *Method 1 (Fixed EKF).* This tracking filter serves as a reference and uses no sophisticated sensor model. The impact of the jammer notch on P_D and the measurement error covariance matrix $\mathbf{R}$ are not taken into account.

2. *Method 2 (Variable EKF).* Here, only the monopulse error covariance $\mathbf{R}(\mathbf{b}_k, \mathbf{j}_k)$ is used as an improvement of the sensor model. The detection probability P_D is is assumed to be constant.

3. *Method 3 (Fixed Pseudo-bearing EKF).* This approach assumes a constant error covariance matrix $\mathbf{R}$, but uses the correct likelihood function, i.e. the jammer notch, in a second-order approximation.

4. *Method 4 (Variable Pseudo-bearing EKF).* In addition to the previous realization, here also the covariance matrix $\mathbf{R}(\mathbf{b}_k, \mathbf{j}_k)$ is part of the sensor model.

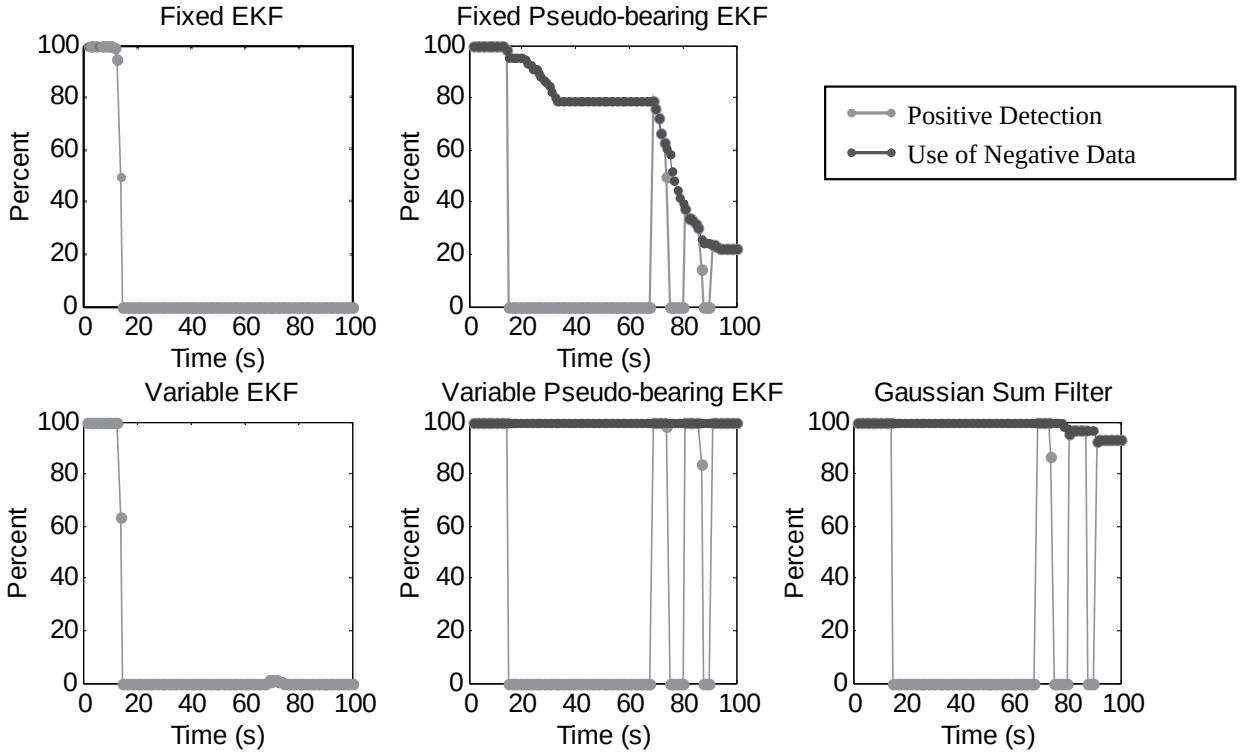


Figure 13: Simulation results (250 runs) characterizing track continuity for different tracking filters.

5. *Method 5 (Gaussian Sum Filter).* In this tracker the complete likelihood function and the monopulse covariance $\mathbf{R}(\mathbf{b}_k, \mathbf{j}_k)$ is used. The number of the mixture components involved to represent $p(\mathbf{x}_k|Z^k)$ is confined by three.

For the methods 3-5 the following is true: If the radar beam points to the vicinity of the blind spot and no detection occurs, a local search is performed. By this, probability mass is concentrated near the blind spot provided the target is actually there.

3.2.3 Selected Simulation Results

Figure 13 shows the mean track continuity averaged over 250 Monte-Carlo runs. The superiority of tracking methods that use context information on the spatial position of the blind spot is obvious. The use of the monopulse covariance matrix is necessary, but not sufficient for avoiding track loss. The methods 3, 4, and 5 can, using "negative" sensor evidence, bridge over the missing data in the jammer notch. In spite of the fact that method 5 is more computationally intensive than method 4, it shows deficiencies if compared with method 4. This is an indication for the fact that further performance improvements are possible by more advanced approximation methods.

A detailed discussion of this approach has been published in:

- W. Blanding, W. Koch , U. Nickel

  Adaptive Phased-Array Tracking in ECM Using Negative Information

  IEEE Transactions on Aerospace and Electronic Systems, vol. 45, nr. 1, p. 152-166, January 2009.

  Abstract

Advances in characterizing the angle measurement covariance for phased array monopulse radar systems that use adaptive beamforming to null a jammer source allow for the use of improved sensor models in tracking algorithms. Using a detection probability likelihood function consisting of a Gaussian sum that incorporates negative contact measurement information, four tracking systems are compared when used to track a maneuvering target passing into and through standoff jammer interference. Each tracker differs in how closely it replicates sensor performance in terms of accuracy of measurement covariance and the use of negative information. Only the tracker that uses both the negative contact information and corrected angle measurement covariance is able to consistently reacquire the target when it exits the jammer interference.

Keywords: Target tracking, adaptive beamforming, standoff jamming, Gaussian sum filter.

3.3 Negative Sensor Information

More advanced sensor models especially enable the exploitation of 'negative' sensor evidence. By this we mean the rigorous drawing of conclusions from expected but actually missing sensor measurements. These conclusions aim at an improvement of the position or velocity estimates for objects currently kept under track. Even a failed attempt to detect an object in the field of view of a sensor is to be considered as a useful sensor output, which can be processed by using appropriate sensor models, i.e. by background information on the sensors, with benefits for target tracking, sensor management, and sensor data fusion. The technical term chosen here for denoting such pieces of evidence, i.e. 'negative' information, seems to be accepted in the data fusion community (see, e.g. [21, 22]).

3.3.1 A Ubiquitous Notion

A very simple example illustrates that negative sensor information is an ubiquitous phenomenon, which often appears in disguise. The notion fits well into the Bayesian formalism. Assume a sensor producing at discrete time instants t_k mutually independent measurements $\mathbf{z}_k$ of a single object with Gaußian likelihood $\mathcal{N}(\mathbf{z}_k;\, \mathbf{H}\mathbf{x}_k,\, \mathbf{R})$. Absence of clutter is assumed ($\rho_F = 0$). The objects are detected with a constant detection probability $P_\mathrm{D} < 1$. We thus have classical Kalman filtering under the constraint that there exists not at each time a measurement. The likelihood function thus yields:

1. In the case of a positive sensor output ($m_k = 1$), $\mathbf{z}_k$ is processed by Kalman filtering leading to $p(\mathbf{x}_k|\mathcal{Z}^k) = \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k}\big)$ with $\mathbf{x}_{k|k}$ and $\mathbf{P}_{k|k}$ given by:

$$\mathbf{P}_{k|k} = (\mathbf{P}_{k|k-1}^{-1} + \mathbf{H}^\top \mathbf{R}^{-1}\mathbf{H})^{-1} \tag{27}$$

$$\mathbf{x}_{k|k} = \mathbf{P}_{k|k}(\mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k-1} + \mathbf{H}^\top \mathbf{R}^{-1}\mathbf{z}_k). \tag{28}$$

2. For a negative sensor output ($m_k = 0$), the likelihood function is a constant $1 - P_\mathrm{D}$. By filtering the prediction density is not modified: $\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1}$, $\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1}$. According to 27 and 28 this result could formally be interpreted as the processing of a fictitious measurement with an infinite measurement error covariance $\mathbf{R}$, since $\mathbf{R}^{-1} = 0$.

3.3.2 Lessons Learned from Examples

The Bayes formalism and the likelihood function thus precisely indicate, in which way a negative sensor output, i.e. a missing detection has to be processed. This observation notion can be generalized and leads to the following conclusions:

1. Missing but expected (i.e. negative) sensor data can convey information on the current target position or a more abstract function of the kinematic object state. This type of negative evidence can be included

in data fusion within the rigorous Bayesian structure. There is no need for recourse to ad hoc or empirical schemes.

2. The prerequisite for processing negative evidence is a refined sensor model, which provides additional background information for explaining its data. As a consequence, negative evidence often appears as an artificial sensor measurement, characterized by a corresponding measurement matrix and a measurement error covariance.

3. The particular form of the fictitious measurement equation involved is determined by the underlying model of the sensor performance, while the fictitious measurement error covariance is characterized by sensor parameters such as sensor resolution, radar beam width, or minimum detectable velocity.

4. Negative evidence implies well-defined probability densities of the object states that prove to be Gaussian mixtures with potentially negative coefficients summing up to one. Intuitively speaking, these components reflect that the targets keep a certain distance from each other, from the last beam position, or the clutter/jammer notch.

5. If the fictitious measurement depends on the underlying sensor-to-target geometry, we can even introduce the fusion of negative evidence.

A detailed discussion of this approach has been published in:

- W. Koch

  On exploiting 'negative' sensor evidence for target tracking and sensor data fusion

  International Journal on Information Fusion, Volume 8, Issue 1, p.28-39, Elsevier, January 2007 (Special Issue: "Best Papers of FUSION 2004", Eds: P. Svensson, J. Schubert, invited paper).

  Abstract
  In various applications of target tracking and sensor data fusion all available information related to the sensor systems used and the underlying scenario should be exploited for improving the tracking/fusion results. Besides the individual sensor measurements themselves, this especially includes the use of more refined models for describing the sensor performance. By incorporating this type of background information into the processing chain, it is possible to exploit 'negative' sensor evidence. The notion of 'negative' sensor evidence covers the conclusions to be drawn from expected but actually missing sensor measurements for improving the position or velocity estimates of targets under track. Even a failed attempt to detect a target is a useful sensor output, which can be exploited by appropriate sensor models providing background information. The basic idea is illustrated by selected examples taken from more advanced tracking and sensor data fusion applications such as group target tracking, tracking with agile beam radar, ground-moving target tracking, or tracking under jamming conditions.

  Keywords: Negative information/evidence, target tracking, sensor resolution, local search, adaptive beam positioning, GMTI sensor fusion

4 Integration of Advanced Object Properties

In several applications, it is necessary to learn more from the sensor data received than the time-varying geolocation of moving objects of interest. Rather, we wish to understand *what* the objects we observe are, i.e. we aim to learn as much as possible about their attributes in order to be able to classify or even identify them. Many relevant object attributes can be derived even from their purely kinematic properties such as speed, heading vector, and normal acceleration as well as from mutual interrelations inferable from multiple object tracks, as has been extensively discussed in the introductory section.

This illustrates this concept with selected application examples and shows how object attributes such as the spatial extension of an extended object or a collectively moving group, including size, shape, and orientation, or even the "anomalous" behavior of an individual in a person stream can be considered as state quantities that can be included into a more general notion of an object state and be tracked by fusing imperfect data within the Bayesian framework.

In particular, the notion of an 'object extension' is introduced by symmetrical and positively definite random matrices serving as state quantities that complement the kinematic state vectors. In this way, matrix-variate analysis is brought into play, by which is is made possible to deal with collectively moving object groups and extended objects in a unified approach. This point of view is all the more appropriate, the smaller the mutual distances between the individual objects within a group are, or the larger an extended object is.

In another example, chemical sensors are discussed that make it possible to classify objects with respect to characteristic chemical signatures. Due to their fundamental lack of space-time resolution, chemical sensors develop their full potential for the classification of individuals only if the output of multiple chemical sensors distributed in space is fused with kinematic person or object tracks. The fusion result enables to identify which individual person in a person stream, for example, is actually carrying a hazardous carry-on item. Obviously, this type of behavior is a fairly well defined pattern of "anomalous behavior" that can easily be recognized using methodologies of multiple sensor data fusion.

4.1 Extended Object Tracking

Due to the increasing resolution capabilities of modern sensors, there is an increasing need for recognizing extended objects as individual units, for initiating extended object tracks, and for extended object track maintenance. Extended objects typically involve a relatively large and often strongly fluctuating number of sensor reports originated by the individual scattering centers that are part of one and the same object. In this context, we usually cannot assume that in subsequent target illuminations the same scattering centers are always responsible for the measurements. The individual sensor reports can therefore no longer be treated in analogy to point object measurements produced by a group of well-separated targets.

Related problems arise if a group of closely-spaced objects is to be tracked. For sensors such as radar, the resolution capability in range is usually much better than in cross-range. As a consequence, two or more targets within the group can be irresolvable, depending on the current sensor-to-target geometry [14, 15, 23]. In addition, little is known about the measurement error of irresolved measurements produced by an unknown number of targets involved. Practically important examples are aircraft formations or ground moving convoys. Under these circumstances, it seems to be reasonable to treat the group as an individual object and to estimate and track its current extension from the sensor data.

The object extension should be considered as an additional 'internal degree of freedom' characterizing an extended object or a collectively moving object group (cluster) to be tracked. The object extension is thus a part of the object state and has to be estimated jointly with the kinematic properties involved. This paper section discusses a realization of this concept within a Bayesian framework. Temporally changing object extensions are tractable within the proposed framework. An extension increasing along a certain direction, e.g., can indicate that an object is beginning to separate into individual subgroups or parts, which then have to be tracked individually.

4.1.1 Generalized Formalism

In a Bayesian view, a tracking algorithm for an extended object or a collectively moving object group is an updating scheme for $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$ at each time t_k given the accumulated sensor data $\mathcal{Z}^k = \{Z_l, m_l\}_{l=1}^k$ and underlying models describing the object's temporal evolution and the sensor performance. Evidently the joint density

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k) = p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^k)\, p(\mathbf{X}_k|\mathcal{Z}^k) \tag{29}$$

can be written as a product of a vector-variate probability density $p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^k)$ and a matrix-variate density $p(\mathbf{X}_k|\mathcal{Z}^k)$ [24]. Furthermore, the probabilistic formalism indicates that the density $p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^k)$, describing the kinematical object properties in the product representation in Eq. 29, should show an explicit dependency on the current object extension $\mathbf{X}_k$. To the author's knowledge, random matrices were first introduced for describing physical phenomena by Eugene Wigner [26].

Extended target tracking, i.e. the iterative calculation of the joint density $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$, basically consists of two steps: prediction and filtering. This scheme is completed by the notion of retrodiction.

Prediction Each update of the joint probability density $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$ of the extended target state $(\mathbf{x}_k, \mathbf{X}_k)$ is preceded by a *prediction* step,

$$p(\mathbf{x}_{k-1}, \mathbf{X}_{k-1}|\mathcal{Z}^{k-1}) \xrightarrow[\text{models}]{\text{evolution}} p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}), \tag{30}$$

based on the underlying evolution models. More precisely, we interpret the prediction density $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1})$ as a marginal density to be calculated by integration:

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}) = \int d\mathbf{x}_{k-1} d\mathbf{X}_{k-1} \; p(\mathbf{x}_k, \mathbf{X}_k|\mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) \; p(\mathbf{x}_{k-1}, \mathbf{X}_{k-1}|\mathcal{Z}^{k-1}). \tag{31}$$

For the (joint) transition density in the previous representation,

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) = p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1})\, p(\mathbf{X}_k|\mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}), \tag{32}$$

we make use of natural Markov-type assumptions for its kinematical part, i.e. $p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) = p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})$, and assume that the object's kinematical properties have no impact on the temporal evolution of the object extension and previous measurements if $\mathbf{X}_{k-1}$ is given, i.e.:

$$p(\mathbf{X}_k|\mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) = p(\mathbf{X}_k|\mathbf{X}_{k-1}). \tag{33}$$

This restriction can be justified in many practical cases. We thus have:

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathbf{x}_{k-1}, \mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) = p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})\, p(\mathbf{X}_k|\mathbf{X}_{k-1}). \tag{34}$$

The probabilistic formalism clearly indicates that the evolution of the object kinematics, described by $p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})$, is affected by the current object extension $\mathbf{X}_k$ as well. This dependence cannot be ignored.

With the previous filtering update $p(\mathbf{x}_{k-1}, \mathbf{X}_{k-1}|\mathcal{Z}^{k-1})$ we obtain the following prediction formula:

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}) = \int & d\mathbf{x}_{k-1} d\mathbf{X}_{k-1} \\ & \times \underbrace{p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})\, p(\mathbf{X}_k|\mathbf{X}_{k-1})}_{\text{evolution model}} \underbrace{p(\mathbf{x}_{k-1}|\mathbf{X}_{k-1}, \mathcal{Z}^{k-1})\, p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1})}_{\text{previous update}}. \end{aligned} \tag{35}$$

The transition densities $p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})$ and $p(\mathbf{X}_k|\mathbf{X}_{k-1})$ will be specified in Section III using suitable models that describe the temporal evolution of extended or group targets.

Further discussion is much simplified if we additionally assume that the *temporal change* of the object extension has no impact on the prediction of the *kinematical* object properties, i.e. if we are allowed to assume $p(\mathbf{x}_{k-1}|\mathbf{X}_{k-1}, \mathcal{Z}^{k-1}) \approx p(\mathbf{x}_{k-1}|\mathbf{X}_k, \mathcal{Z}^{k-1})$ or, in other words, to replace $\mathbf{X}_{k-1}$ by $\mathbf{X}_k$. Such an assumption seems to be justified in many practical cases. By this approximation, the predicted density

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}) = p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^{k-1})\, p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \tag{36}$$

is given by two factors to be obtained by independent integrations:

$$p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{X}_k, \mathbf{x}_{k-1})\, p(\mathbf{x}_{k-1}|\mathbf{X}_k, \mathcal{Z}^{k-1})\, d\mathbf{x}_{k-1} \tag{37}$$

$$p(\mathbf{X}_k|\mathcal{Z}^{k-1}) = \int p(\mathbf{X}_k|\mathbf{X}_{k-1})\, p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1})\, d\mathbf{X}_{k-1}. \tag{38}$$

Filtering The prediction is followed by a *filtering* step, in which the current sensor information Z_k at time t_k is to be processed:

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}) \xrightarrow[\text{sensor model}]{\text{data: } Z_k, n_k} p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k). \tag{39}$$

More precisely, in the filtering step the sensor-specific likelihood function $p(Z_k, n_k|\mathbf{x}_k, \mathbf{X}_k)$, defined by the current data and the underlying sensor model, is combined with the predicted density by exploiting Bayes' formula [27, 9]:

$$p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k) = \frac{p(Z_k, n_k|\mathbf{x}_k, \mathbf{X}_k)\, p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1})}{\int p(Z_k, n_k|\mathbf{x}_k, \mathbf{X}_k)\, p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1})\, d\mathbf{x}_k d\mathbf{X}_k}. \tag{40}$$

4.1.2 Extended Object Prediction

The probability density $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$ of an extended or group target state is given by Eq. 40. The joint densities in this equation can be written as products:

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k) &= p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^k)\, p(\mathbf{X}_k|\mathcal{Z}^k) \\ p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^{k-1}) &= p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^{k-1})\, p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \\ p(Z_k, n_k|\mathbf{x}_k, \mathbf{X}_k) &= p(Z_k|n_k, \mathbf{x}_k, \mathbf{X}_k)\, p(n_k|\mathbf{x}_k, \mathbf{X}_k). \end{aligned} \tag{41}$$

The kinematical state variable $\mathbf{x}_k$ at time t_k is given by $\mathbf{x}_k = (\mathbf{r}_k^\top, \dot{\mathbf{r}}_k^\top, \ddot{\mathbf{r}}_k^\top)^\top$ with the spatial state components $\mathbf{r}_k$. Let the dimension d of the vector $\mathbf{r}_k$ be also the dimension of the $d \times d$ SPD matrix $\mathbf{X}_k$ that describe the current ellipsoidal object extension (SPD: symmetrical and positively definite). $\dot{\mathbf{r}}_k$, $\ddot{\mathbf{r}}_k$ denote the corresponding velocity and acceleration. The dimension of the kinematical state vector $\mathbf{x}_k$ is thus $s \times d$, where $s - 1$ describes up to which derivative the object kinematics is modeled. Here we have $s = 3$.

Extended Object Evolution The temporal evolution of an extended or collective object is modeled as usual in Kalman filtering theory: $\mathbf{x}_k = \mathbf{\Phi}_{k|k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$, $p(\mathbf{v}_k) = \mathcal{N}(\mathbf{v}_k;\, \mathbf{0},\, \mathbf{\Delta}_{k|k-1})$. Using the Kronecker product [24], the evolution matrix $\mathbf{\Phi}_{k|k-1}$ can be written as:

$$\mathbf{\Phi}_{k|k-1} = \mathbf{F}_{k|k-1} \otimes \mathbf{1}_d, \tag{42}$$

where the $s \times s$ matrix $\mathbf{F}_{k|k-1}$ is given for example by van Keuk's or Singer's model. The use of Kronecker products will prove to be very convenient in the subsequent calculations. For the dynamics noise covariance $\mathbf{\Delta}_{k|k-1}$, we postulate the following structure:

$$\mathbf{\Delta}_{k|k-1} = \mathbf{D}_{k|k-1} \otimes \mathbf{X}_k. \tag{43}$$

Model parameters describing the underlying dynamics are part of a $s \times s$ matrix $\mathbf{D}_{k|k-1}$, as given by van Keuk's model, for example. The $s \times s$ matrices $\mathbf{F}_{k|k-1}$, $\mathbf{D}_{k|k-1}$ also appear in this form in the 1D tracking problem. The system noise is thus a band limited Gaussian acceleration noise process with a covariance proportional to the extension matrix $\mathbf{X}_k$. This has the effect of directing the acceleration of the group (or object) centroid along the direction of the major axis of the ellipse.

The assumption of a dynamics covariance matrix $\mathbf{\Delta}_{k|k-1}$ depending on the current object extension $\mathbf{X}_k$, which is a consequence of the probability formalism, needs a discussion with more physical arguments:

1. The collective character of a group motion is the more pronounced the smaller the group is. The dynamical behavior of a smaller group is thus to a larger extent deterministic in nature ("'maneuvering becomes dangerous"').

2. For a group dissolving into subgroups, i.e. if its extension is increasing, the knowledge of its dynamical behavior decreases, and the motion of the group is not easily predictable, being expressed by the increasing dynamics noise covariance.

3. In addition, large extended or group objects will produce so many sensor measurements that the prediction part of the tracking process, i.e. exploitation of information on the object evolution, seems to be negligible if compared to the gain obtained in the filtering step.

4. In case of extended objects like submarines or ground moving convoys, which show a clear orientation, the proposed dynamics model provides a natural description of their actual movement along the major axes of the extension ellipse.

Besides these more or less physically motivated reasons, an important formal argument exists in favor of the model: A dynamics model of the proposed form implies a formal structure of the densities $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$, which enables a rigorous application of the Bayesian formalism under certain assumptions.

Structure of the Predicted Density According to Eq. 41, the kinematics can be discussed separately from the extension estimation in the tracking process. Let us assume that the density of the kinematical state variable $p(\mathbf{x}_{k-1}|\mathbf{X}_k, \mathcal{Z}^{k-1})$ after filtering at time t_{k-1} is a Gaussian with the following special structure:

$$p(\mathbf{x}_{k-1}|\mathbf{X}_k, \mathcal{Z}^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1};\, \mathbf{x}_{k-1|k-1},\, \mathbf{P}_{k-1|k-1} \otimes \mathbf{X}_k). \tag{44}$$

Then the previous evolution model guarantees that this structure is preserved by the prediction process (Eq. 37):

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{X}_k, \mathcal{Z}^{k-1}) &= \int \mathcal{N}(\mathbf{x}_k;\, (\mathbf{F}_{k|k-1} \otimes \mathbf{1}_d)\mathbf{x}_{k-1},\, \mathbf{D}_{k|k-1} \otimes \mathbf{X}_k) \\ &\qquad \times\ \mathcal{N}(\mathbf{x}_{k-1};\, \mathbf{x}_{k-1|k-1},\, \mathbf{P}_{k-1|k-1} \otimes \mathbf{X}_k)\, d\mathbf{x}_{k-1} \qquad (45) \\ &= \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k-1},\, \mathbf{P}_{k|k-1} \otimes \mathbf{X}_k) \qquad (46) \end{aligned}$$

according to the usual rules for Kronecker products with $\mathbf{x}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ given by:

$$\mathbf{x}_{k|k-1} = (\mathbf{F}_{k|k-1} \otimes \mathbf{1}_d)\mathbf{x}_{k-1|k-1} \tag{47}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k|k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k|k-1}^{\top} + \mathbf{D}_{k|k-1} \tag{48}$$

in close analogy to standard Kalman filtering.

Moreover, let us assume that the densities of the extension state variable $p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1})$ are given by Inverted Wishart densities [24] defined up to a factor independent of $\mathbf{X}_{k-1}$ by:

$$\begin{aligned} p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1}) &= \mathcal{IW}\big(\mathbf{X}_{k-1};\, \nu_{k-1|k-1},\, \mathbf{X}_{k-1|k-1}\big) \qquad (49) \\ &\propto |\mathbf{X}_{k-1}|^{-\frac{1}{2}\nu_{k-1|k-1}}\ \mathrm{etr}\big[-\tfrac{1}{2}\mathbf{X}_{k-1|k-1}\mathbf{X}_{k-1}^{-1}\big]\,. \qquad (50) \end{aligned}$$

d is the dimension of the measurement vectors $\mathbf{z}_k^j$ and $\mathrm{etr}[\mathbf{A}]$ an abbreviation for $\exp[\mathrm{tr}\mathbf{A}]$ with $\mathrm{tr}\mathbf{A}$ denoting the trace of a matrix $\mathbf{A}$. The expectation of $\mathbf{X}_{k-1}$ is given by $\mathbb{E}[\mathbf{X}_{k-1}] = \frac{\mathbf{X}_{k-1|k-1}}{\nu_{k-1|k-1}-2d-2}$.

In the prediction step, the parameters $\nu_{k|k-1}$, $\mathbf{X}_{k|k-1}$ defining $p(\mathbf{X}_k|\mathcal{Z}^{k-1})$ have to be calculated from $\nu_{k-1|k-1}$, $\mathbf{X}_{k-1|k-1}$ available after the previous filtering step according to appropriate modeling assumptions. In a first heuristic approach, we postulate that the expectation of the predicted density shall be equal to the expectation of the previous filtering step; i.e.: $\frac{\mathbf{X}_{k|k-1}}{\nu_{k|k-1}-2d-2} = \frac{\mathbf{X}_{k-1|k-1}}{\nu_{k-1|k-1}-2d-2}$. The degrees of freedom of an inverse Wishart density are related to the 'precision' of the corresponding expectation. The 'precision' of predictions, however, will decrease with increasing update intervals $\Delta t_k = t_k - t_{k-1}$. With a temporal

decay constant τ as an additional modeling parameter, the following prediction update equations seem to be plausible:

$$\nu_{k|k-1} = \mathrm{e}^{-\Delta t_k/\tau}\, \nu_{k-1|k-1} \tag{51}$$

$$\mathbf{X}_{k|k-1} = \frac{\mathrm{e}^{-\Delta t_k/\tau}\nu_{k-1|k-1}-d-1}{\nu_{k-1|k-1}-d-1}\, \mathbf{X}_{k-1|k-1}. \tag{52}$$

$\tau = \infty$ represents a static object or group extension.

4.1.3 Extended Object Filtering

In the case of extended or group targets, the significance of a single measurement is obviously dominated by the underlying object extension. The sensor-specific measurement error that describe the precision by which a given scattering center is currently measured is the more unimportant, the larger the actual extension of the object is compared to the measurement error. The individual measurements must therefore rather be interpreted as measurements of the centroid of the extended or collective object, since it is unimportant for the extended object tracking task which of the varying scattering centers was actually responsible for the measurement.

We thus interpret each individual measurement produced by an extended object as a measurement of the object centroid with a corresponding 'measurement error' that is proportional to the object extension $\mathbf{X}_k$ to be estimated. By means of this 'measurement error', however, the object extension $\mathbf{X}_k$ becomes explicitly part of the likelihood function $p(Z_k, n_k|\mathbf{x}_k, \mathbf{X}_k)$, which describes what the measured quantities Z_k, n_k can say about the state variables $\mathbf{x}_k$ and $\mathbf{X}_k$. As a consequence of this interpretation, the object extension $\mathbf{X}_k$ can also be estimated by exploiting the sensor data (besides the kinematical state vector $\mathbf{x}_k$).

By using the Kronecker product, we also assume that the measurement matrix has the following special structure:

$$(h_k^1\mathbf{1}_d, h_k^2\mathbf{1}_d, h_k^3\mathbf{1}_d) = \mathbf{H}_k \otimes \mathbf{1}_d. \tag{53}$$

With $\mathbf{H}_k = (1, 0)$, e.g., scenarios with range and azimuth measurements are accessible after transforming them into Cartesian coordinates. According to the previous considerations, the corresponding measurement error covariance is given by the extension matrix $\mathbf{X}_k$ to be estimated.

Likelihood Function In order to exploit Bayes' formula Eq. 40, a likelihood function factorized according to Eq. 41 needs to be defined. For the sake of simplicity, let us exclude false or unwanted measurements at present. In a first approximation, the number n_k of measurements in Z_k is assumed to be independent of the state variables $\mathbf{x}_k$, $\mathbf{X}_k$; i.e. $p(n_k|\mathbf{x}_k, \mathbf{X}_k)$ is assumed to be a constant. The joint density $p(Z_k|m_k, \mathbf{x}_k, \mathbf{X}_k)$ can be factorized:

$$p(Z_k|m_k, \mathbf{x}_k, \mathbf{X}_k) \propto \mathcal{N}\big(\mathbf{z}_k;\, (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_k,\, \tfrac{\mathbf{X}_k}{m_k}\big)\ \mathcal{LW}\big(\mathbf{Z}_k;\, m_k-1,\, \mathbf{X}_k\big). \tag{54}$$

with a centroid measurement $\mathbf{z}_k$, a corresponding scattering matrix $\mathbf{Z}_k$, and a Wishart density in $\mathbf{Z}_k$ with m_k-1 degrees of freedom.

Structure after Filtering With these preliminaries, it is possible to exploit the Bayes formula Eq. 40. To this end, we have to calculate the product:

$$\begin{aligned} p(Z_k|n_k, \mathbf{x}_k, \mathbf{X}_k)\, p(\mathbf{x}_k, \mathbf{X}_k|\mathbf{Z}^{k-1}) \propto\ & \mathcal{N}\big(\mathbf{z}_k;\, (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_k,\, \tfrac{\mathbf{X}_k}{n_k}\big) \\ & \times\ \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{k|k-1},\, \mathbf{P}_{k|k-1} \otimes \mathbf{X}_k\big) \\ & \times\ \mathcal{LW}\big(\mathbf{Z}_k;\, n_k-1,\, \mathbf{X}_k\big)\ \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k-1},\, \mathbf{X}_{k|k-1}\big). \end{aligned} \tag{55}$$

By standard calculations (product formula for Gaussians and properties of Kronecker products, the product of the two Gaussians in the previous equation yields:

$$\mathcal{N}\big(\mathbf{z}_k;\, (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_k,\, \tfrac{\mathbf{X}_k}{n_k}\big)\; \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{k|k-1},\, \mathbf{P}_{k|k-1} \otimes \mathbf{X}_k\big) = \mathcal{N}\big(\mathbf{z}_k;\, (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_{k|k-1},\, S_{k|k-1}\mathbf{X}_k\big)\; \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k} \otimes \mathbf{X}_k\big) \quad (56)$$

where the quantities $\mathbf{x}_{k|k}$ and $\mathbf{P}_{k|k}$ are given by

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + (\mathbf{W}_{k|k-1} \otimes \mathbf{1}_d)(\mathbf{z}_k - (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_{k|k-1}) \quad (57)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{W}_{k|k-1} S_{k|k-1} \mathbf{W}_{k|k-1}^\top \quad (58)$$

with a scalar *innovation factor* and a gain matrix defined by

$$S_{k|k-1} = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \tfrac{1}{n_k} \quad (59)$$

$$\mathbf{W}_{k|k-1} = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top S_{k|k-1}^{-1}. \quad (60)$$

The first factor on the right side in Eq. 54 does not depend on the kinematical state variable $\mathbf{x}_k$. It can be rewritten as

$$\mathcal{N}\big(\mathbf{z}_k;\, (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_{k|k-1},\, S_{k|k-1}\mathbf{X}_k\big) \propto |\mathbf{X}_k|^{-\frac{1}{2}}\; \mathrm{etr}\big[-\tfrac{1}{2}\mathbf{N}_{k|k-1}\mathbf{X}_k^{-1}\big] \quad (61)$$

up to a factor independent of the state variables and with an *innovation matrix* $\mathbf{N}_{k|k-1}$ defined by

$$\mathbf{N}_{k|k-1} = S_{k|k-1}^{-1}\big(\mathbf{z}_k - (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_{k|k-1}\big)\,\big(\mathbf{z}_k - (\mathbf{H}_k \otimes \mathbf{1}_d)\mathbf{x}_{k|k-1}\big)^\top. \quad (62)$$

The remaining two factors on the right side of Eq. 55 yield:

$$\mathcal{LW}\big(\mathbf{Z}_k;\, n_k - 1,\, \mathbf{X}_k\big)\, \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k-1},\, \mathbf{X}_{k|k-1}\big)\, |\mathbf{X}_k|^{-\frac{1}{2}}\; \mathrm{etr}\big[-\tfrac{1}{2}\mathbf{N}_{k|k-1}\mathbf{X}_k^{-1}\big] \propto \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k},\, \mathbf{X}_{k|k}\big) \quad (63)$$

with the simple update equations:

$$\mathbf{X}_{k|k} = \mathbf{X}_{k|k-1} + \mathbf{N}_{k|k-1} + \mathbf{Z}_k \quad (64)$$

$$\nu_{k|k} = \nu_{k|k-1} + n_k. \quad (65)$$

Joint Density after Filtering The probability density function of the joint state $(\mathbf{x}_k, \mathbf{X}_k)$ after processing the current sensor data Z_k at time t_k is thus given by:

$$p(\mathbf{x}_k, \mathbf{X}_k | \mathcal{Z}^k) = \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k} \otimes \mathbf{X}_k)\; \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k},\, \mathbf{X}_{k|k}\big). \quad (66)$$

Important Remark: By means of the *innovation matrix* $\mathbf{N}_{k|k-1}$, it is possible to estimate an unknown measurement error covariance even in the case of point source targets or the extension of a completely irresolved target group, i.e. for $n_k = 1$.

4.1.4 Extended Object Kinematics

In many practical applications, we are interested in estimates of the kinematic state variables only, i.e. on the marginal density $p(\mathbf{x}_k|\mathcal{Z}^k)$ obtained by integrating the joint density $p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)$ over the random matrices $\mathbf{X}_k$:

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \int p(\mathbf{x}_k, \mathbf{X}_k|\mathcal{Z}^k)\, d\mathbf{X}_k \quad (67)$$

$$= \int \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k} \otimes \mathbf{X}_k)\; \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k},\, \mathbf{X}_{k|k}\big)\, d\mathbf{X}_k. \quad (68)$$

By lengthy but elementary algebraic calculations, the integrand can be transformed into the following product:

$$\mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{k|k},\, \mathbf{P}_{k|k} \otimes \mathbf{X}_k\big)\; \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k},\, \mathbf{X}_{k|k}\big) \propto \\ |\mathbf{Y}_k(\mathbf{x}_k)|^{-\frac{(\nu_{k|k}+s-sd)+sd}{2}}\; \mathcal{IW}\big(\mathbf{X}_k;\, \nu_{k|k}+s,\, \mathbf{Y}_k(\mathbf{x}_k)\mathbf{X}_{k|k}\big) \quad (69)$$

with a matrix $\mathbf{Y}_k = \mathbf{Y}_k(\mathbf{x}_k)$ depending on the kinematical state variable $\mathbf{x}_k$ whose determinant is given by

$$|\mathbf{Y}_k| = 1 + (\mathbf{x}_k - \mathbf{x}_{k|k})^\top\, (\mathbf{P}_{k|k}^{-1} \otimes \mathbf{X}_{k|k}^{-1})(\mathbf{x}_k - \mathbf{x}_{k|k}). \quad (70)$$

With this representation of the integrand, integration over the random matrix $\mathbf{X}_k$ is trivial. We ultimately find that the marginal density with respect to the kinematical state variable $\mathbf{x}_k$ is given by a multivariate version of the Student density with $\nu_{k|k}$ degrees of freedom:

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \mathcal{T}\big(\mathbf{x}_k;\, \nu_{k|k}+s-sd,\, \mathbf{x}_{k|k}, \mathbf{P}_{k|k} \otimes \mathbf{X}_{k|k}\big). \quad (71)$$

By exploiting the multivariate t-density a 'gating' can be constructed that is simply a version of the Hotelling-t^2-test.

It is immediately clear that the marginalized prediction and retrodiction densities are also given by Student densities: $p(\mathbf{x}_l|\mathcal{Z}^{l-1}) = \mathcal{T}\big(\mathbf{x}_l;\, \nu_{l|l-1}+s-sd,\, \mathbf{x}_{l|l-1}, \mathbf{P}_{l|l-1} \otimes \mathbf{X}_{l|l-1}\big)$, $p(\mathbf{x}_l|\mathcal{Z}^k) = \mathcal{T}\big(\mathbf{x}_l;\, \nu_{l|k}+s-sd,\, \mathbf{x}_{l|k}, \mathbf{P}_{l|k} \otimes \mathbf{X}_{l|k}\big)$.

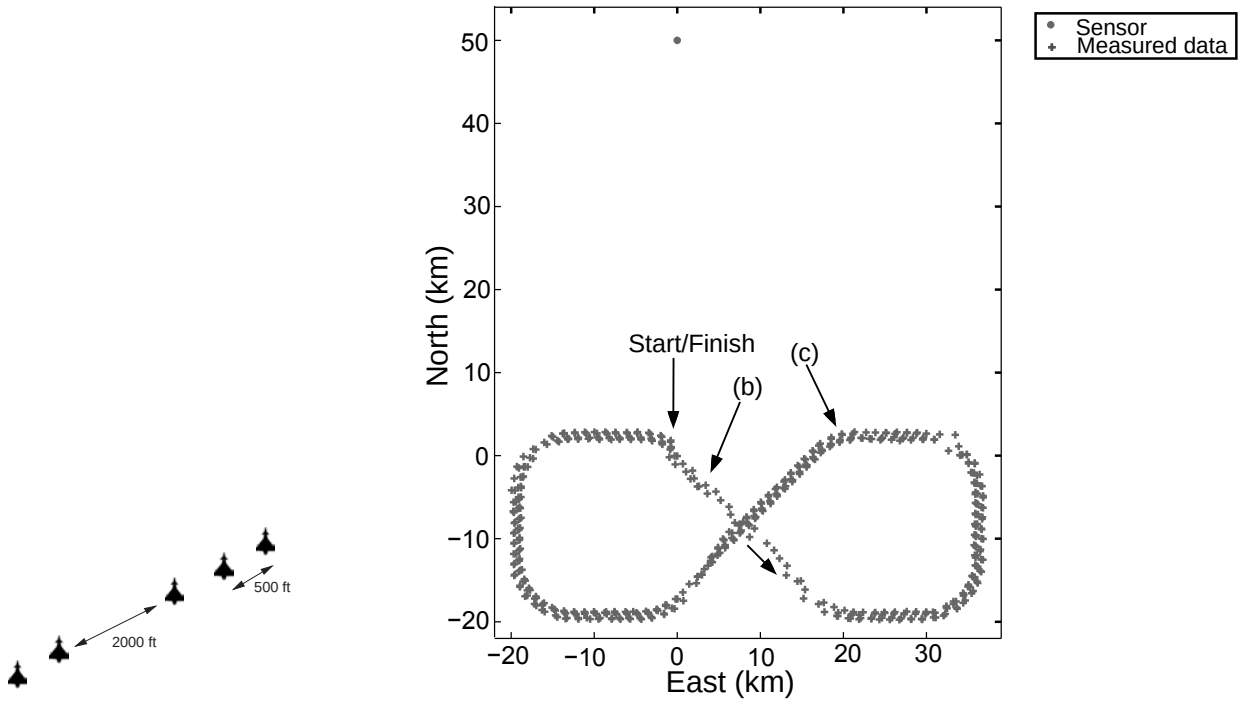
4.1.5 Selected Simulation Results

For the sake of simplicity, aircraft trajectories are simulated in a plane and partitioned into straight and circular segments where each aircraft is moving with a constant tangential speed as shown in Figure 14. In an echelon formation consisting of five aircraft, the leading aircraft is responsible for navigating, while the other aircraft try to preserve their relative position to the leading aircraft. The underlying radar sensor has a finite resolution capability (range resolution: 50 m, azimuth resolution: $1.0°$). The corresponding measurement error standard deviations for resolvable objects are 10 m and $0.1°$, respectively. The orientation of the aircraft formation varies as it moves around the trajectory. The update interval is 5 s. For the parameters of the Van-Keuk-evolution model, we chose Σ =1 g, θ = 40 s. The normal acceleration during the maneuvers is 1 g, the speed is 250 m/s. The formation starts at the origin of the coordinate system.

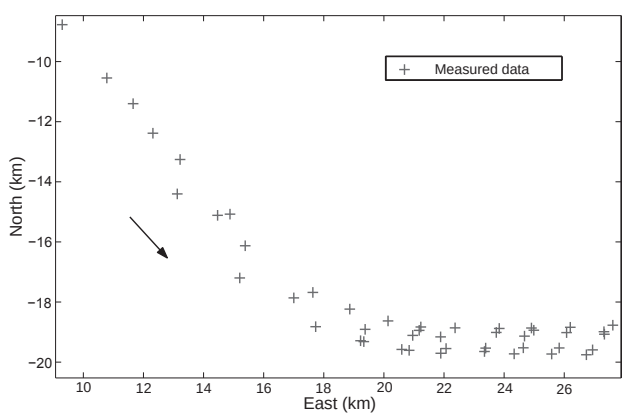
Simulating a Partly Resolvable Formation For the simulation of radar measurements, the corresponding measurements errors and the sensor resolution have to be taken into account. The generation of false returns is not considered here. For a group of two targets at positions (r_1, φ_1), (r_2, φ_2) in polar coordinates with respect to the sensor position, the probability of being unresolved, P_u, can be modeled:

$$P_u(\Delta r, \Delta\varphi) = \mathrm{e}^{-\frac{1}{2}(\Delta r/\alpha_r)^2}\, \mathrm{e}^{-\frac{1}{2}(\Delta\varphi/\alpha_\varphi)^2} \quad (72)$$

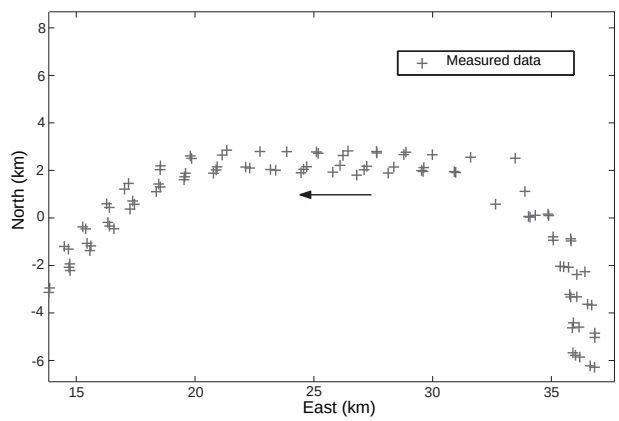
with $\Delta r = r_2 - r_1$, $\Delta\varphi = \varphi_2 - \varphi_1$, where the sensor parameters α_r, α_φ characterize the radar's resolution capability in range and azimuth, respectively. According to this probability and for pairs of aircraft, it can be simulated whether an unresolved measurement occurs or not. In case of a resolution conflict the pair is replaced by a single unresolvable object at the centroid position. For large formations with more than two targets, a list is created containing all possible pairs of aircraft. A pair of this list is selected at random according to P_u and merged. In this case, one of the aircraft is to be removed the list, which thus has to be recalculated. If no resolution conflict occurs according to the probability $1-P_u$, the pair is removed from the list. The previous reasoning is repeated for the remaining pairs. If the list is empty, the algorithm terminates.



(a) Accumulated radar data.



Detail (b)



Detail (c)

Figure 14: Measurements of a Partly Unresolved Formation (resolution: 50 m, 1.0°, measurement error: 10 m, 0.1°)

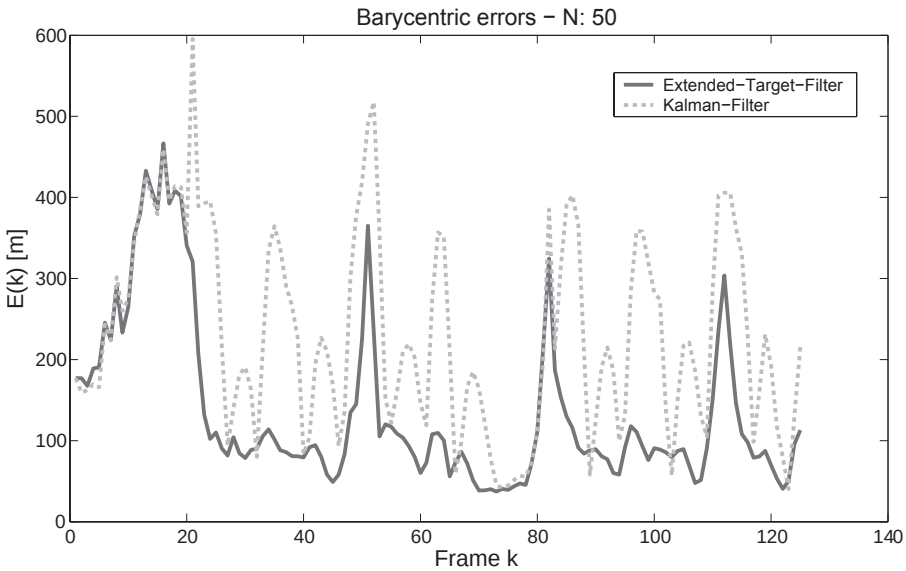


Figure 15: Position error: extended target filter vs. standard Kalman filter

We finally have to consider the effect of successive mergings on the simulated measurement errors of unresolvable objects. To this end, we assume that an unresolved measurement error resulting from m aircraft is to be simulated according to $\sigma^u_{r,\varphi} = m\sigma_{r,\varphi}$ where $\sigma_{r,\varphi}$ denote the standard deviations of resolvable range and azimuth measurements, respectively. It is reasonable to delimit the growth of the measurement error by the sensor resolution: $\sigma^u_{r,\varphi} \leq \alpha_{r,\varphi}$. In the same manner, missing detections can be simulated. We here assumed a detection probability $P^u_D = 1$ for unresolvable aircraft and $P^r_D = 0.9$ otherwise.

Impact of the resolution parameters Figure 14 displays the radar data simulated according to these assumptions. The details in Figures 14b, c clearly reveal the impact of resolution phenomena and make it obvious that they depend heavily on the current sensor-to-target geometry. The discussed phenomena make it clearly evident that even a very regular target formation is very similar in appearance to an extended object producing a highly fluctuating number of measurements. There is no reasonable hope to be able to track the single components of such a formation individually.

Discussion of results As before in the case of a totally irresolvable formation, in Figure 15 the root mean squared errors of the position estimates of the extended target filter are compared with the corresponding results produced by standard Kalman filtering. As the measurement error in the Kalman filter, we used the scattering matrix calculated from the true target positions within the formation and processed averaged measurements. The extended target filter shows significantly smaller estimation errors.

In Figure 16, the estimated major semi-axes are compared with the major semi-axes of the scattering matrix of the true target positions. The concordance seems to be fairly good. The peak in the middle of the time axis is due to the reorganization of the formation.

Figures 17 and 18 a ‘split-off’ maneuver that is clearly indicated by the increasing eccentricity of the estimated extension ellipse. As soon as the eccentricity exceeds a certain threshold, two extended object tracks are initiated and the sub-groups are tracked separately. The proposed filter thus provides a criterion of when a single extended object track has to be split into two extended object tracks. An analogous mechanism is possible in the case of a larger formation being created by merging two or more converging sub-groups.

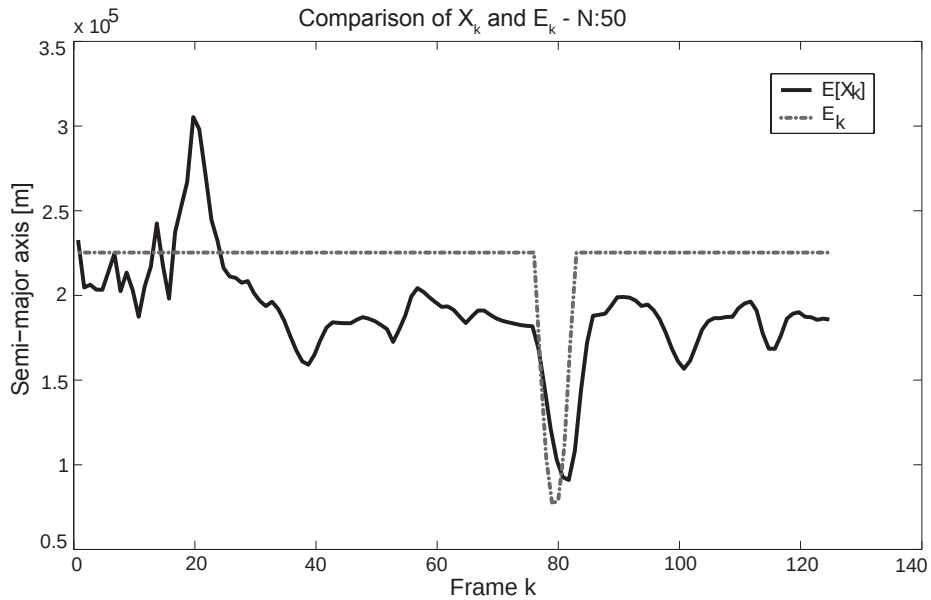


Figure 16: Mean major semi-axes of $\mathbb{E}[\mathbf{X}_k|\mathcal{Z}^k]$ vs. $\mathbf{X}_k$

4.1.6 Summary of Results

The essential theoretical result of this paper seems to be the insight that the Bayesian formalism can be applied to extended objects or collectively moving target clusters with approximations to be justified in many applications. Basically, the application of the Bayesian formalism relies on closure properties of matrix-variate Wishart and Inverted Wishart densities under multiplication.

In view of practical applications, the following aspects seem to be of particular relevance:

1. There exists a natural extension of the standard Kalman filter equations to objects whose spatial extension is approximately described by ellipsoids.

2. The object extension can be modeled by symmetrical, positive definite random matrices, whose statistical properties are described by well-known matrix-variate probability densities [24].

3. Due to the mild character of the approximations used, a representation of the probability densities involved by particle filtering techniques such as proposed in [27], does not seem to be necessary. The densities are characterized by a finite parameter set.

4. Information on the objects' kinematic properties is represented by vector-variate Student densities. 'Gating', i.e. exclusion of unwanted measirements, is provided by a Hotelling test.

5. Tracking of point source targets with an unknown measurement error is a limiting case of the proposed method (e.g. tracking of an irresolvable formation).

6. With respect to the kinematical properties, the achievable filter performance is only slightly different from Kalman filtering with a known measurement error covariance matrix.

7. The estimated measurement error covariance matrix corresponds to the true measurement error covariance matrix (simulated) relatively well. This is an interesting side result, considering the small number of data in the case of a totally unresolved group.

8. The proposed filter can successfully be applied to target formations, which are only partly resolvable depending on the underlying sensor-to-target geometry.

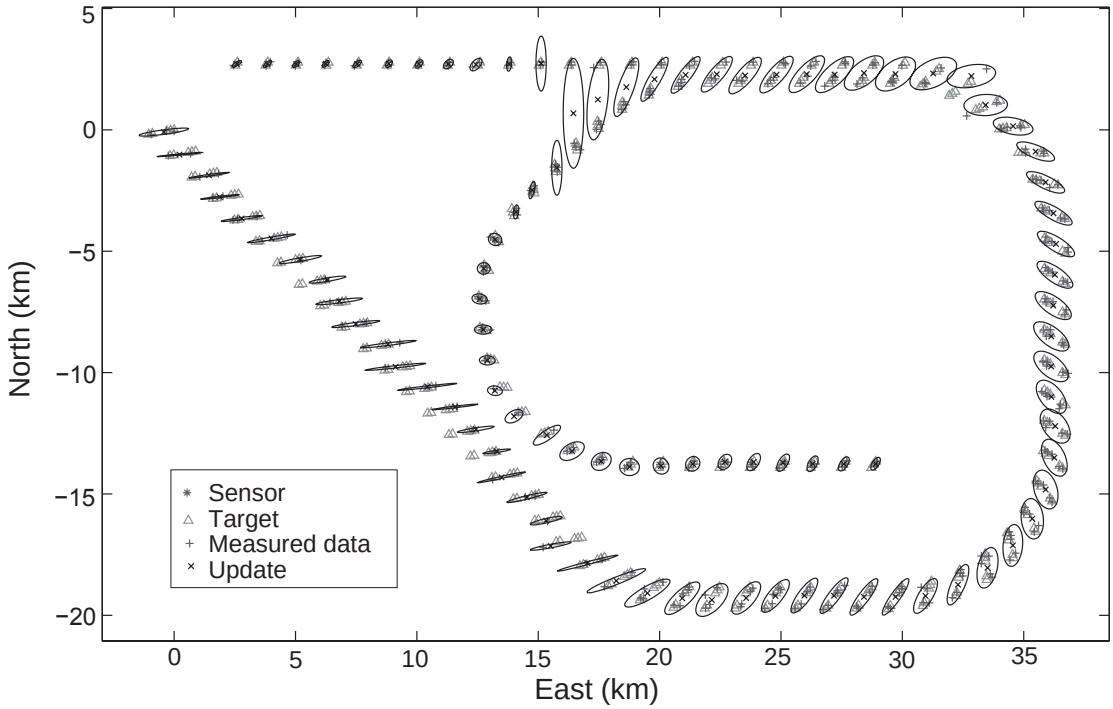


Figure 17: Echelon formation: split-off maneuver

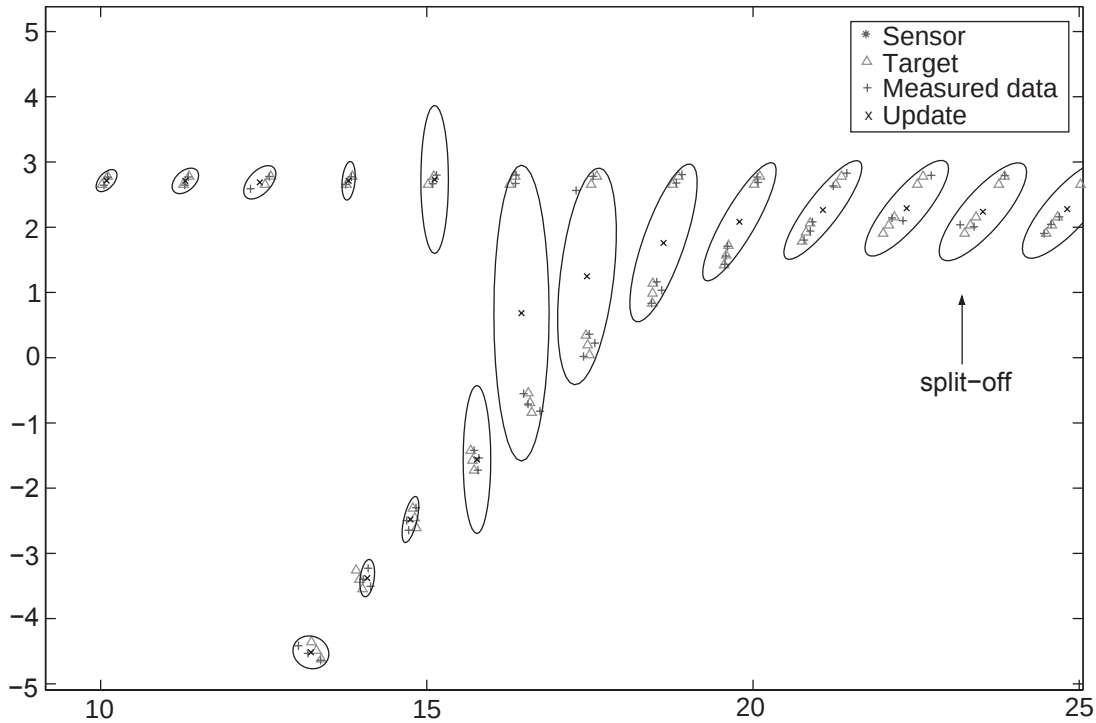


Figure 18: Split-off maneuver (detail)

9. "Split-off" maneuvers, indicating that an object is beginning to separate into individual subgroups or parts, can be detected by analyzing the extension ellipsoid (e.g. by designing a test based on its eccentricity).

In principle, the proposed approximate Bayesian method for dealing with extended objects or collectively moving target clusters can be embedded into multiple-object, multiple-hypothesis tracking techniques and can also be combined with context information (e.g. road-map assisted convoy tracking). This opens an interesting field for further research.

A detailed discussion of this approach has been published in:

- W. Koch

  Bayesian Approach to Extended Object and Cluster Tracking using Random Matrices.

  IEEE Transactions on Aerospace and Electronic Systems, Vol. 44, Nr. 3, p. 1042-1059, July 2008.

  Abstract
  In algorithms for tracking and sensor data fusion, the targets to be observed are usually considered as point source objects; i.e. compared to the sensor resolution, their extension is neglected. Due to the increasing resolution capabilities of modern sensors, however, this assumption is often no longer valid, since different scattering centers of an object can cause distinct detections when passing the signal processing chain. Examples of extended targets are found in short-range applications (littoral surveillance, autonomous weapons, or robotics). A collectively moving target group can also be considered as an extended target. This point of view is the more appropriate, the smaller the mutual distances between the individual targets are. Due to the resulting data association and resolution conflicts, any attempt to track the individual objects within the group seems to be no longer reasonable.

  With simulated sensor data produced by a partly unresolvable aircraft formation, the addressed phenomena are illustrated, and an approximate Bayesian solution to the resulting tracking problem is proposed. Ellipsoidal object extensions are modeled by random matrices, which are treated as additional state variables to be estimated or tracked. We expect that the resulting tracking algorithms are also relevant for tracking large, collectively moving target swarms.

  Keywords: Target tracking, extended targets, group targets, target clusters, sensor resolution, random matrices, matrix-variate analysis

4.2 Classification with Chemical Sensors

This deficiency, however, can be compensated by fusing the output of multiple chemical sensors distributed in space with kinematic data produced by laser-range-scanners or video cameras. In other words, tracking spans an additional temporal dimension for processing chemical sensor attributes. Multiple person tracking and chemical attribute fusion are thus to be performed within a single framework (see figure 4.2, [28, 29]).

In designing a multiple sensor system for decision support in security applications we wish to know which person going through an access area in an airport, e.g., may be carrying explosives. With reference to the experimental corridor sketched in figure 2, five chemical sensors measure at each scan the chemical signatures with respect to the position of each of the chemical sensors symbolized (green filled circles). Furthermore, there two laser-range-scanners (cyan and blue filled rectangles) are used as tracking sensors. In this example, three persons are walking along the corridor, one of them carrying hazardous material. Space-time processing for multiple person tracking and classification obviously plays a key role: Only in an integrated framework can the potential of chemical sensors or other attribute sensors of this type be exploited entirely.

A detailed discussion of this approach has been published in:

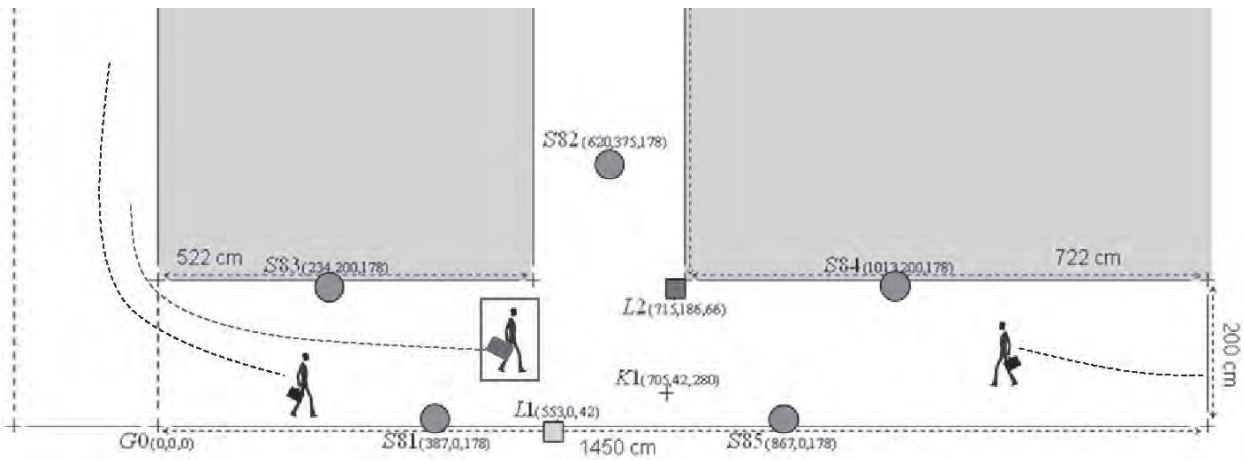


Figure 19: Experimental corridor with 5 chemical and 2 tracking sensors.

- M. Wieneke and W. Koch

  Combined person tracking and classification in a network of chemical sensors

  Elsevier International Journal of Critical Infrastructure Protection, vol. 2, nr. 1-2, p. 51-67, May 2009.

  Abstract
  Transportation infrastructures play a crucial role in the way of life and the economic vitality of a modern society. Access points like stations, airports or harbors are among the most critical elements in these infrastructures because they offer a possibility to bring in hazardous materials that can be used for attacks against people and against the transportation network itself. A timely recognition of such threats is essential and can be significantly supported by systems that monitor critical areas continuously and call the security personnel in case of anomalies. We are describing the concept and the realization of an indoor security assistance system for real time decision support. The system is specifically designed for the surveillance of entrance areas in transportation access facilities and consists of multiple heterogeneous sensors: Chemical sensors detecting hazardous materials provide data for the classification of persons. But due to their limited spatial■temporal resolution, a single chemical sensor cannot localize a substance and assign it to a person. We compensate for this deficiency by fusing the output of multiple, distributed chemical sensors with kinematical data from laser range■scanners. Both tracking and fusion of tracks with chemical attributes can be processed within a single framework called Probabilistic Multi-Hypothesis Tracking (PMHT). An extension of PMHT for dealing with classification measurements (PMHT-c) already exists.We show how PMHT- c can be applied to assign chemical attributes to person tracks. This affords the localization of threats and a timely notification of the security personnel.

  Keywords: Person tracking, Probabilistic Multiple Hypothesis Tracking (PMHT), Classification, Attributes, Data fusion, Security assistance systems

References

[1] Martin E. Liggins, David L. Hall, and James Llinas (Eds.). *Handbook of Multisensor Data Fusion – Theory and Practice*. CRC Press, 2nd Edition (2008).

[2] E. Bosse, J. Roy, and S. Wark. *Concepts, Models, and Tool for Information Fusion*. Artech House (2007).

[3] Belur V. Dasarathy (Ed.). "A Special Issue on information fusion in computer security". In *Information Fusion*, vol. 10, nr. 4 (2009).

[4] Michael Wunder and Jürgen Grosche (Eds.). *Verteilte Führungsinformationssysteme*. Springer (2009).

[5] Ludger Schmidt, Christopher M. Schlick, and Jürgen Grosche (Eds.). *Ergonomie und Mensch-Maschine-Systeme*. Springer (2008).

[6] Sam S. Blackman. "Multiple Hypothesis Methods". In *IEEE AES Magazine*, **19**, 1, p. 41–52, 2004.

[7] Sam S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems.* Artech House (1999).

[8] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation.* Wiley & Sons (2001).

[9] W. Koch. "Target Tracking". Chapter 8 in Stergios Stergiopoulos (Ed.). *Advanced Signal Processing: Theory and Implementation for Sonar, Radar, and Non-Invasive Medical Diagnostic Systems*. CRC Press (2001).

[10] O. E. Drummond. "Target Tracking with Retrodicted Discrete Probabilities". In *Proceedings of SPIE **3163**, Signal and Data Processing of Small Targets*, 249 (1997).

[11] W. Koch. "Fixed-Interval Retrodiction Approach to Bayesian IMM-MHT for Maneuvering Multiple Targets". In *IEEE Transactions on Aerospace and Electronic Systems*, AES-36, No. 1, (2000).

[12] W. Koch, J. Biermann. "Anomaly Detection from Tracking-derived Situation Elements". In *Processdings of the NATO Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness - MSA 2009*, La Spezia, Italy, September 2009.

[13] B. Ristic, B. La Scala, M. Morelande, N. Gordon. Statistical Analysis of Motion Patterns in AIS Data: Anomaly Detection and Motion Prediction', In *Proc. of the 11th International Conference on Information Fusion*, FUSION 2008, July 2008, Cologne.

[14] F.E. Daum and R.J. Fitzgerald. "The Importance of Resolution in Multiple Target Tracking". In *SPIE Signal and Data Processing of Small Targets*, Vol. 2235, p. 329 (1994).

[15] W. Koch and G. van Keuk. "Multiple Hypothesis Track Maintenance with Possibly Unresolved Measurements". In *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 33, No. 3 (1997).

[16] W. Koch. "On Exploiting 'Negative' Sensor Evidence for object Tracking and Sensor Data Fusion". In *Information Fusion Journal*, **8**(1), 28-39, 2007.

[17] W. Koch. "Experimental Results on Bayesian MHT for Maneuvering Closely-Spaced Objects in a Densely Cluttered Environment". In *IEE International Radar Conference RADAR'97*, p. 729, Edinburgh, UK (1999).

[18] W. Koch. "On Bayesian MHT for Formations with Possibly Unresolved Measurements – Quantitative Results". In *SPIE Signal and Data Processing of Small Targets*, Vol. 3163, p. 417, (1997).

[19] NICKEL, U., Performance Measure of Corrected Adaptive Monopulse Estimation. In: *IEE Proceeding on Radar, Sonar, and Navigation*, 146, pp. 17-24, Feb. 1999.

[20] W. Blanding, W. Koch, and U. Nickel. "Adaptive Phased-Array Tracking in ECM Using Negative Information". Accepted for publication in IEEE Transactions on Aerospace and Electronic Systems, expected publication date January 2009. Pereliminary version appeared in Proc. FUSION 2006, Florence (2006).

[21] AGATE, K.J.S., 'Utilizing Negative Information to Track Ground Vehicles Through Move-stop-move Cycles', in: Proceedings of *Signal Processing, Sensor Fusion, and Target Recognition XIII*, **SPIE Vol. 5429**, Orlando, FL, April 2004.

[22] SIDENBLADH, H., 'Multi-Target Particle Filtering for the Probability Hypothesis Density', in: Proceedings of *The Seventh International Conference on Information Fusion (FUSION 2003)*, pp. 800-806, Cairns, Australia, 2003.

[23] Blom, Henk A.P., Bloem, Edwin A., "Approximate Bayesian tracking of two targets that maneuver in and out formation amidst unresolved and false measurements". To appear in: *IEEE Transactions on Transaction in Aerospace and Electronic Systems* (2008).

[24] Gupta, A. K., Nagar, D.K., *Matrix Variate Distributions.* Chapman & Hall/CRC, 1999.

[25] Harville, D. A., *Matrix Algebra from a Statistician's Perspective*. Springer, 1997.

[26] Wigner, E. P., Random matrices in physics. In: SIAM Review, 9, 1-13, 1967.

[27] Branko Ristic, Sanjeev Arulampalam, and Neil Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House Radar Library (2004).

[28] M. Wieneke and W. Koch. "Combined person tracking and classification in a network of chemical sensors". *International Journal of Critical Infrastructure Protection*, In Press, Accepted Manuscript, Available online 24 November 2008, Elsevier.

[29] Becher, C., Foresti, G.L., Kaula, P., Koch, W., Lorenz, F.P., Lubczyk, D., Micheloni, C., Piciarelli, C., Safenreiter, K., Siering, C., Varela, M., Waldvogel, S., Wieneke, M. "A Security Assistance System combining Person Tracking with Chemical Attributes and Video Event Analysis". In *Proc. of 11th ISIF International Conference on Information Fusion*, Cologne, July 2008